Instrumenting the Musician

Measuring and Enhancing
Affective and Behavioural Interaction
During Collaborative Music Making

Evan Lloyd Morgan

Submitted in partial fulfilment
of the requirements of the
Degree of Doctor of Philosophy

School of Electronic Engineering and Computer Science

Queen Mary University of London

February 2016

Statement of Originality

I, Evan Lloyd Morgan, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: 

Date: 01/02/2016

For details of collaboration and publications see Section 1.5.

Instrumenting the Musician

Measuring and Enhancing Affective and Behavioural Interaction During Collaborative Music Making

Evan Lloyd Morgan

Abstract

Modern sensor technologies facilitate the measurement and interpretation of human affective and behavioural signals, and have consequently become widely used tools in the fields of affective computing, social signal processing and psychophysiology. This thesis investigates the use and development of these tools for measuring and enhancing affective and behavioural interaction during collaborative music making.

Drawing upon work in the aforementioned fields, an exploratory study is designed, where self-report and continuous behavioural and physiological measures are collected from pairs of improvising percussionists. The findings lead to the selection of gaze, motion, and cardiac activity as input measures in the design of a device to enhance affective and behavioural interaction between co-present musicians. The device provides musicians with real-time visual feedback on the glances or body motions of their co-performers, whilst also recording cardiac activity as a potential measure of musical decision making processes. Quantitative evidence is found for the effects of this device on the communicative behaviours of collaborating musicians during an experiment designed to test the device in a controlled environment. This study also reports findings on discrete and time series relationships between cardiac activity and musical decision-making. A further, qualitative study is designed to evaluate the appropriation and impact of the device during long-term use in naturalistic settings. The results provide insights into earlier findings and contribute towards an empirical understanding of affective and behavioural interaction during collaborative music making, as well as implications for the design and deployment of sensor-based technologies to enhance such interactions.

This thesis advances the dominant single-user paradigm within human-computer interaction and affective computing research, towards multi-user scenarios, where the concern is human-human interaction. It achieves this by focusing on the emotionally rich, and under-studied context of co-present musical collaboration; contributing new methods and findings that pave the way for further research and real-world applications.

Acknowledgements

I would like to thank my primary supervisor, Dr. Hatice Gunes, for her guidance and attentive supervision over the last four years. It has been a pleasure, and I wish you the best of luck in your new position at the University of Cambridge. Thanks also to my secondary supervisor, Dr. Nick Bryan-Kinns, who provided a valuable input to the work in this thesis, especially on the musical interaction front.

During my time at QMUL numerous academics and members of staff have provided support and guidance. In particular, I would like to acknowledge and thank Dr. Marcus Pearce, for his advice concerning the processing of EEG data and use of the IDyOM software; Prof. Pat Healey, for general advice and encouragement; Prof. Geraint Wiggins, for providing guidance as the external assessor in my stage reviews; Richard Kelly, for his all-round support as the MAT programme manager; Kok Ho Huen, for always having a solution to my technical quandaries; and Geetha Bommireddy, for her assistance with lab facilities and equipment.

My friends and fellow students on the MAT programme are a fantastic group of multi-talented people, who have provided inspiration and support over the last four years. It has been a pleasure to get to know you all, and I hope we stay in touch. Thanks also to my friends outside of QMUL, who've provided fun and distraction along the way; and especially to Josh, Michael, and Ronnie Rosenberg for their warm hospitality during my final months in London.

My parents, Pat Evans and Kenton Morgan, provide unwavering support, encouragement, and friendship; and I've been especially grateful for that over the last few months. Thanks!

Finally, I'd like to thank Lauren Clay. Despite being all the way up in Edinburgh, you've been an ever-present source of smiles, encouragement, and musical expertise; and I'm looking forward to our adventures ahead!

This work was funded by the Engineering and Physical Sciences Research Council (EPSRC) as part of the Centre for Doctoral Training in Media and Arts Technology at Queen Mary University of London (ref: EP/G03723X/1).

Contents

List of Figures 11

List of Tables 14

1 Introduction 18

1.1 Motivation 19
1.2 Aims and Approach 21
1.3 Thesis Structure 22
1.4 Contributions 23
1.4.1 New Methods Proposed 24
1.4.2 Findings Obtained 25
1.5 Associated Publications 26

2 Background 28

2.1 Measuring Behaviour and Affect 28
2.1.1 Relevant Disciplines 28
2.1.2 Cognitive Measurement 31
2.1.3 Behavioural Measurement 32
2.1.4 Physiological Measurement 33
2.2 Co-Present Collaborative Music Making 34
2.2.1 Group Interactions 34
2.2.2 Group Musical Improvisation 36
2.2.3 Studying Co-present Musical Interactions 38
2.3 Summary 39

3 Study 1: Sensing Collaborating Musicians 40

3.1 Aims and Motivation 40
3.2 Research Design and Data Collection 41

3.2.1 Participants . . . 42
3.2.2 Measures . . . 42
3.2.3 Data Synchronisation . . . 44
3.2.4 Setup . . . 44
3.2.5 Tasks . . . 45
3.3 Data Processing . . . 45
3.3.1 Preparation . . . 45
3.3.2 Feature Extraction . . . 46
3.4 Analyses and Results . . . 49
3.4.1 Individual-Level Analyses . . . 49
3.4.2 Dyad-Level Analyses . . . 52
3.4.3 Study-Level Analyses . . . 55
3.5 Discussion . . . 60
3.5.1 Design . . . 60
3.5.2 Findings . . . 62
3.5.3 Measures . . . 66
3.6 Summary . . . 68

4 The LuminUs: Designing a Device for Collaborating Musicians 69
4.1 Design Considerations . . . 69
4.1.1 Input Measures . . . 70
4.1.2 Output Modalities . . . 71
4.2 Related Work . . . 71
4.2.1 Gaze . . . 72
4.2.2 Motion . . . 77
4.2.3 Cardiac Activity . . . 82
4.2.4 Visual Feedback . . . 90
4.2.5 Summary and Relevant Findings . . . 94
4.3 The LuminUs: Overview . . . 96
4.3.1 Sensors and Inputs . . . 96
4.3.2 Signal Processing . . . 97
4.3.3 Feedback Visualisation . . . 98
4.4 The LuminUs: Hardware and Software . . . 99
4.4.1 Hardware . . . 99
4.4.2 Software . . . 101
4.5 Discussion . . . 110

4.6 Summary 111

5 Study 2: Testing the LuminUs in the Lab 112

5.1 Aims 113

5.1.1 On the Effects of the LuminUs 113

5.1.2 On Cardiac Activity and Musical Decision Making 114

5.2 Research Design and Data Collection 114

5.2.1 Animation 115

5.2.2 Participants 117

5.2.3 Setup 117

5.2.4 Tasks 120

5.2.5 Measures 121

5.2.6 Software 122

5.3 Data Processing 123

5.3.1 Gaze Features 124

5.3.2 ECG Features 124

5.3.3 Accelerometer Features 125

5.3.4 Decision Making Features 125

5.3.5 LuminUs Feedback Features 131

5.4 Analyses and Results: Effects of the LuminUs Feedback 131

5.4.1 Gaze Behaviour 131

5.4.2 Body Motion 137

5.4.3 Musical Decision Making 140

5.4.4 Self-report Measures 144

5.5 Analyses and Results: Cardiac Activity and Musical Decision Making . 147

5.5.1 Time-averaged Analyses 147

5.5.2 Time Series Analyses 151

5.6 Discussion 164

5.6.1 Effects of the LuminUs Feedback 165

5.6.2 Cardiac Activity and Musical Decision Making 169

5.6.3 Experimental Design 174

5.7 Conclusions 176

5.8 Summary 178

6 Study 3: Investigating the LuminUs in Practice 180

6.1 Aims 181

6.2 Review of Qualitative Research Methods 183

6.2.1 Data Collection Methods 184
6.2.2 Analysis Methods 185
6.3 Research Design and Data Collection 187
6.3.1 Software and Hardware 187
6.3.2 Participants 190
6.3.3 Sessions 191
6.3.4 Data Collection 192
6.4 Thematic Analysis 194
6.4.1 Method 194
6.4.2 Data 196
6.4.3 Coding 196
6.4.4 Theme Development 197
6.5 Final Themes 201
6.5.1 Expression and Communication 201
6.5.2 Group Attributes 208
6.5.3 Technology 210
6.5.4 Aspects of Music 223
6.5.5 Context 224
6.6 Discussion 225
6.6.1 Appropriation 225
6.6.2 Meaning 227
6.6.3 Impact 229
6.6.4 Design 229
6.7 Summary 232

7 Discussion and Implications **233**
7.1 Sensing Musical Collaboration 234
7.1.1 Sensing the Musician 234
7.1.2 Gaze 236
7.1.3 Motion 238
7.1.4 Musical Output 239
7.1.5 Physiology 240
7.2 Collection and Analysis of Dyadic Data 242
7.2.1 Study Designs 242
7.2.2 Data Collection Methods 244
7.2.3 Analysis Methods 246

7.3 Design Implications . . . 249
7.3.1 Potential Applications . . . 249
7.3.2 Sensor Selection . . . 251
7.3.3 Data Processing . . . 252
7.3.4 Visual Feedback . . . 253
7.4 Summary . . . 254

8 Conclusion 256
8.1 Overview and Major Findings . . . 256
8.1.1 Study 1: Sensing Collaborating Musicians . . . 256
8.1.2 The LuminUs: Designing a Device for Collaborating Musicians . 257
8.1.3 Study 2: Testing the LuminUs in the Lab . . . 258
8.1.4 Study 3: The LuminUs in Practice . . . 259
8.2 Contributions . . . 260
8.2.1 New Methods Proposed . . . 260
8.2.2 Findings Obtained . . . 262
8.3 Limitations . . . 263
8.3.1 Study Groups . . . 263
8.3.2 Statistical Analyses . . . 264
8.3.3 Sensor Technologies . . . 265
8.3.4 External and Ecological Validity . . . 266
8.4 Future Work . . . 266
8.4.1 Taking the LuminUs into the Wild . . . 266
8.4.2 Giving the LuminUs Intelligence . . . 267
8.4.3 Utilising Wearable and Everyday Sensors . . . 268
8.4.4 Designing Affective and Intelligent Musical Instruments . . . 268
8.4.5 Assisting Musicians with Disabilities . . . 269
8.5 Closing Remarks . . . 269

Bibliography 270

Appendices 307

A Study 1 Materials 308
A.1 Ethical Approval . . . 309
A.2 Participant Forms . . . 310
A.3 Strip Plot Visualisations . . . 315

A.4 SPSS Linear Mixed Model Command Syntaxes 318

B Study 2 Materials **319**

B.1 Ethical Approval 320

B.2 Participant Forms 321

B.3 Post-performance Survey 324

B.4 Time-series Analyses Plots 329

C Study 3 Materials **333**

C.1 Ethical Approval 334

C.2 Participant Forms 335

C.3 Thematic Analysis Codes 338

C.4 Follow-up Survey 341

List of Figures

2.1 Schematic of the dual pathway to creativity model. 37

3.1 A still image of the captured video for Study 1. 43
3.2 Image illustrating the setup and the equipment used in Study 1. 43
3.3 The post performance questionnaire that was used in the study. 44
3.4 A short segment of MIDI data, illustrating how rhythmic change points were identified. 48
3.5 Strip-plot of individual participant data for a two window segment of a single improvisation task. 50
3.6 Plot of a segment of heart rate data showing detected extrema and corresponding timing of rhythmic change points. 50
3.7 Plots of the mean heart rates during the 7 seconds before and after rhythmic change points (RCPs) for participants with significant RCP-extrema alignments. 54
3.8 Plot of the mean heart rate for all participants during the 7 seconds before, and after all the Rhythmic Change Points from both conditions. 58

4.1 A review of audio level displays on contemporary audio devices. 90
4.2 The 'Icebreaker' T-shirt 92
4.3 Images of the devices used to collect data from collaborating musicians. 97
4.4 Early design sketch for the LuminUs 99
4.5 The NeoPixel stick, used for the light output of the LuminUs. 100
4.6 The Pro Micro microcontoller, used to communicate with and control the lights on the LuminUs. 100
4.7 Diagram showing the connections between the two NeoPixel LED sticks and the Pro Micro microcontroller. 101
4.8 Model views of the assembled LuminUs casing. 102
4.9 Exploded model of the LuminUs casing. 103

4.10 The software graphical user interface created for the Shimmer sensors. . 105
4.11 Images of the Pupil software interface. 106
4.12 The circular marker used for gaze detection. 106
4.13 The LuminUs with all LEDs lit in gaze feedback mode, and motion feedback mode. 108
4.14 A system diagram, showing the five software programs that are used to operate the LuminUs. 109

5.1 Film strips of the four short-listed animations. 116
5.2 Annotated image of the experimental setup for Study 2. 118
5.3 Block diagram of the experimental setup for Study 2. 119
5.4 The control software graphical user interface for Study 2. 123
5.5 Example plot of change point features and MIDI notes for piano data. . 129
5.6 Plot of the animation frame difference with cue points and still images. 130
5.7 Bar graphs showing the mean number of glances; the mean time spent glancing; and the mean glance duration within each of the seven conditions. 132
5.8 Bar graph showing the mean pianist and percussionist body motion within each of the seven conditions. 138
5.9 Bar graphs showing the mean number of musical changes; and the mean information content within each of the seven conditions. 140
5.10 Stacked bar graph showing the cumulative self-report responses within the conditions where both participants receiving gaze feedback; motion feedback; or no feedback. 145
5.11 Bar graph showing the mean creativity ratings for the percussionists and pianists for each attempt within each of the conditions. 146
5.12 Musical changes and information content averaged over all performances for percussionists and pianists. 152
5.13 Musical changes and information content averaged over all performances. 153
5.14 Detrended heart rate; heart rate variability; and LF/HF ratio averaged over all performances for percussionists and pianists. 154
5.15 Analysis of time series correlations between musical changes and HRV over all performances, for percussionists only. 157
5.16 Analysis of time series correlations between musical changes and HRV over all performances, for pianists only. 158
5.17 Analysis of time series correlations between mean information content and HRV over all performances for percussionists only. 159

5.18 Analysis of time series correlations between mean information content and HRV over all performances, for pianists only. 160
5.19 Plot of the CCFs between the mean HRV and the mean number of changes time series for each participant. 164
5.20 Plot of the CCFs between the mean information content and the mean HRV time series for each participant. 164

6.1 The LuminUs software interface used in Study 3. 188
6.2 Examples of two of the fiducial markers used for gaze detection. 189
6.3 An image showing how one of the participants set up their equipment during the study. 189
6.4 The candidate thematic map. 199
6.5 The final thematic map. 200

7.1 Sensing musical collaborations. 234

B.4.1 Analysis of time series correlations between musical changes and LF/HF ratio over all performances, for percussionists only 329
B.4.2 Analysis of time series correlations between musical changes and LF/HF ratio over all performances for pianists only 330
B.4.3 Analysis of time series correlations between information content and LF/HF ratio over all performances, for percussionists only. 331
B.4.4 Analysis of time series correlations between information content and LF/HF ratio over all performances for pianists only. 332

List of Tables

3.1 A summary of the measures, sensors, and features used in Study 1. . . . 48
3.2 Percentages of rhythmic change points and heart rate extrema that are aligned for individual participants within each condition. 52
3.3 Spearman correlation coefficient values for significant within-dyad correlations between self-report ratings. 53
3.4 Synchrony in continuous heart rate data between participants, within dyads. 55
3.5 Linear mixed effects modelling estimates of fixed effects between continuous features. 57
3.6 Linear mixed effects modelling estimates of fixed effects between self-report measures and continuous features. 59
3.7 Linear mixed effects modelling estimates of fixed effects of visibility condition on all self-report measures and noteworthy continuous measures. . 61

5.1 Details of the four short-listed animations. 115
5.2 Animation survey responses. 117
5.3 Animation cue point descriptions. 131
5.4 Statistical results for the differences between the number of glances in conditions 1-6, relative to condition 7, for participants receiving LuminUs feedback. 133
5.5 Statistical results for the differences between the proportion of glances reciprocated within 5 seconds in conditions 1-6, relative to condition 7. . 135
5.6 Linear mixed effects modelling estimates of fixed effects of two measures of the light feedback provided by the LuminUs upon the number of glances. 136
5.7 Linear mixed effects modelling estimates of cross-participant effects of the number of glances. 136

5.8 Linear mixed effects modelling estimates of fixed effects of two measures of the light feedback provided by the LuminUs upon the body motion of the musicians. 139
5.9 Linear mixed effects modelling estimates of cross-participant effects of the mean body motion. 139
5.10 Statistical results for the differences between the mean information content for percussionists within the conditions where they received LuminUs feedback, relative to the no-feedback condition. 141
5.11 Linear mixed effects modelling estimates of effects of LuminUs feedback upon mean information content and the number of musical change points. 142
5.12 Linear mixed effects modelling estimates of effects of the number of glances and body motion upon the mean information content. 143
5.13 Statistical results for differences between self-report responses across the seven LuminUs feedback conditions. 146
5.14 Linear mixed effects modelling estimates of the effects of self-reported creativity upon various cardiac activity features. 148
5.15 Linear mixed effects modelling estimates of the effects of the number of change points upon various cardiac activity features. 149
5.16 Linear mixed effects modelling estimates of the effects of mean information content upon various cardiac activity features. 150
5.17 Results of cross correlation analysis between HRV and number of changes; and information content. 162

6.1 Group and participant details for Study 3. 190
6.2 An example of one of the interview topic guides used in Study 3. 193
6.3 An example of the coding of an interview transcript. 197

List of Abbreviations

CCF Cross-correlation Function

CP Change Point

ECG Electrocardiogram

EEG Electroencephalography

FFT Fast Fourier Transform

FOV Field Of Vision

GSR Galvanic Skin Response

GUI Graphical User Interface

HCI Human-Computer Interaction

HF High Frequency

HR Heart Rate

HRV Heart Rate Variability

IBI Inter-beat Interval

IC Information Content

IGT Iowa Gambling Task

LED Light Emitting Diode

LF Low Frequency

LMM Linear Mixed Model

MIDI Musical Instrument Digital Interface

NVC Non-verbal Communication

OSC Open Sound Control

PNS Parasympathetic Nervous System

PPQ Post-performance Questionnaire

QoM Quantity Of Motion

RCP Rhythmic Change Point

SCR Skin Conductance Response

SMH Somatic Marker Hypothesis

SNS Sympathetic Nervous System

SR Self-report

SSP Social Signal Processing

SSQS Semi-structured Qualitative Study

TSA Time Series Analysis

Chapter 1

Introduction

In 1997 Rosalind Picard outlined her manifesto for affective computing: "computing that relates to, arises from, or influences emotions" (Picard, 1997, p. 1). At this time the technological landscape was a far cry from the one that confronts us today. Mobile phones were little more than phones without the wires, social networking took place predominantly in the presence of others, and the only creatures tweeting were birds. Huge changes were afoot. These were foreseen, to an impressive extent, by Mark Weiser in his definitions of 'ubiquitous computing' - a term he had coined early in the '90s (Weiser, 1991). Weiser envisaged that computers would become omnipresent, but comparably invisible components in our everyday lives and physical environment. He hypothesised that this change would be driven by the expansion of the Internet, and a human desire "not to be held hostage" by technology (Weiser, 1993). The emergence of affective computing appeared to encapsulate the kind of technological advances that Weiser had in mind. By being able to interact with humans on an emotional level, affective computers could adopt a wider role in our lives, whilst simultaneously appearing less computer-like.

Weiser also understood that the influence of technology on people is not a one-way process. Rather, it is a feedback loop, where new technologies lead to changes in lifestyles, which lead to changes in needs, and consequently investment in new technologies (Weiser, 1999). In the two decades since Weiser made his predictions, technological advances have emerged at such a pace that there has been little opportunity to reflect on the simultaneous changes occurring in the people exposed to these technologies. There is an increasing amount of evidence to suggest that these changes are significant; affecting the way we interact (Drouin and Davis, 2009; Turkle, 2012; Wooldridge and Shapka, 2012; Amichai-Hamburger and Hayat, 2011), and the way we think (Small et al., 2009). Furthermore, it appears that these changes are not merely an extension of the capabilities provided by new technology, they are also derivative of the very fabric on which much of this technology is built - digital data and structured logic. This is most starkly demonstrated in the modern methods that many of us use to represent and express ourselves: the binary 'like' button on Facebook, the length-restricted 'packets' of text sent via Twitter, and our geo-located positions on Google Maps.

In short, while people set out to create technology that was human-centred, it appears that to some extent humans are becoming technology-centred. As our interactions with computers become more commonplace, do we risk undermining some of the evolutionarily complex elements of human interaction? This is a poignant question for researchers in the field of human-computer interaction.

This thesis shifts the focus towards human-human interactions, and the ways in which these interactions could benefit from affect and social signal-sensing technologies. It does this within the context of a particular type of interaction; one which is emotionally rich, universally practised, and uniquely human - **collaborative music making**. Three studies are reported, investigating the use of sensor technologies for measuring affective and behavioural aspects of collaborative music making. Applications for these technologies are also investigated through the design and evaluation of a device that provides performing musicians with sensor-related feedback.

The work in this thesis is informed and inspired by existing research in the fields of affective computing and social signal processing (SSP), which showcase ways in which human affective states and behavioural cues can be sensed, analysed, and categorised using technological and computer-based approaches. Affective computing is typically directed towards single-user, human-computer interaction scenarios, whilst SSP marks a shift towards the consideration of multi-user scenarios. With its focus on collaborative music making, this thesis investigates the measurement of continuous and low-level representations of affect and social interaction during rich human-human interactions. In doing so, it broadens the horizons of affective computing and SSP research, contributing new findings and methodological approaches.

1.1 Motivation

Human-Computer Interaction (HCI) is a vast, multi-faceted, and growing area of research. Its popularity is underscored by the fact that, for many of us, interactions with computers constitute a large proportion of our everyday lives. Sensors play an important role within HCI; serving to mediate interactions between human and machine. Some researchers approach the use of sensors from a computer-oriented perspective;

using them to develop new input modalities such as touch interfaces. For others, sensors are seen as a tool to enable machines to behave more like humans, complimenting our abilities (Wilson, 2012). This latter approach requires the consideration of human factors, such as emotion, behaviour, and social psychology. In particular, sensors have been used to measure and analyse gestures (Glowinski et al., 2011; Ren et al., 2013), body motions (Bianchi-Berthouze, 2012; Sanghvi et al., 2011; Lim et al., 2014), speech (Schuller and Weninger, 2015), facial expressions (Chang et al., 2009; Zeng et al., 2009), gaze (Müller et al., 2012; Vertegaal, 1999), and physiology (Nasoz et al., 2004; Picard et al., 2001). These measures have been used for a variety of purposes, such as detecting engagement (Rich et al., 2010), recognising affective states (Kulic and Croft, 2007), and monitoring mental workload (Fairclough, 2009).

The field of affective computing has embraced the use of sensor technologies in a human-oriented approach to HCI. Research in this field has predominately focused on single-user scenarios, whereby computers are used to elicit and recognise affective and behavioural responses from individuals.(Kory and D'Mello, 2014; Vinciarelli et al., 2012). A founding motivation for the research presented in this thesis was a desire to contribute towards affective computing research by investigating how affective and behavioural sensing could be used to measure and enrich our interactions with other people (Morgan et al., 2013). With respect to the enrichment of interactions, it was implicit that an aspect of this research should focus upon real-world applications for sensor technologies.

Gunes and Schuller (2013) note that selecting a specific context for affective computing research enables the design of systems that are realistic, and simplifies some of the problems associated with automatic analysis and recognition. Given the expanse of human-human interaction situations that occur in everyday life, the decision was made to focus on a particular type of situation. The criteria for selecting this situation were *emotional richness*, *co-presence* (shared space), and the potential for *real-world applications*. Collaborative music making was selected because it satisfies all of these criteria. In particular, it involves cooperation and synchronisation (Reidsma et al., 2014); emotional expression (Biasutti and Frezza, 2009); non-verbal social interaction (Wilson and MacDonald, 2012); and concurrent, rather than sequential contributions (Healey et al., 2005). Indeed, Varni et al. (2008, p. 1) note that "performing arts, and in particular music, are an ideal test-bed to investigate non-verbal expressive communication and expressive gesture".

With respect to the potential for real-world applications, the decision to investigate collaborative music making was motivated by the belief that affect and behaviour-sensing technologies hold great potential for enhancing our experience of the performing arts; where emotional expression is exaggerated and the value of physical presence is highly regarded. Furthermore, the music making community have already demonstrated a willingness to adopt sensor technologies. These are being used in the design of new interfaces for musical expression, which can be controlled through touch, gesture, motion, and physiology (Kuhara and Kobayashi, 2011; Mitchell et al., 2012; Medeiros and Wanderley, 2014). However, despite their heightened ability to sense human input, these devices predominantly focus on single-user operation and do very little to sense and enhance the emotional, interpersonal, and communicative elements of collaborative music making (Carlile and Hartmann, 2005; Fencott and Bryan-Kinns, 2010).

In addition to the motivations above, investigating the use of sensor technologies in the context of collaborative music making was seen as an opportunity to gain a better understanding of the behavioural and affective processes that accompany collaborative music making. By incorporating influences from a variety of fields – including HCI, affective computing, and music interaction – it was also hoped that this thesis would showcase a novel, interdisciplinary approach, and provide inspiration for future work.

1.2 Aims and Approach

The following specific aims were established:

A1 Investigate how best to continuously measure behaviour and affect during dyadic musical interactions: Various sensors can be used for measuring and interpreting human behaviour and affect. These include physiological, cognitive, and behavioural measures. The use of these measures in the context of musical collaboration has received little attention. As such, this research aims to identify and develop specific sensors and techniques that are well suited to the study of musical interactions.

A2 Design and evaluate a sensor-based device for enhancing affective and behavioural interaction during collaborative music making: A core motivation of this research is to contribute towards the realisation of real-world applications for affect and behavioural sensing technologies. Consequently, an aim of this research is to develop and test a device that utilises such technologies in order to enhance collaborative music making.

A3 Provide insights into the wider applications for affect and behaviour sensing technologies in the context of collaborative music making: In addition to evaluating the device described in **A2**, this research aims to identify wider challenges, considerations, and recommendations concerning the design of affect and behaviour sensing devices for collaborating musicians.

An empirical approach is taken to addressing these aims, involving a combination of quantitative and qualitative research methods. The studies in this thesis predominantly investigate improvisation as a form of group music making that necessitates collaboration (Sawyer, 2003). Group improvisation is also an inherently social activity, effected through complex forms of non-verbal communication (Sawyer, 2003; Wilson and MacDonald, 2012); this makes it well suited to the study of affective and behavioural aspects of human interactions. To further constrain the focus of the research, study groups were restricted to dyads.

The first study in this thesis takes an exploratory approach, owing to a lack of existing work on affect and behaviour sensing during collaborative music making. Exploratory studies provide a valid means of generating novel hypotheses from previously unexplored situations (Jaeger and Halliday, 1998). In this case, relationships are explored between experimentally obtained sensor data and aspects of collaborative music making. The findings facilitate the identification of specific sensors, measures, and features, which subsequently inform the design and development of a prototype device. Specific exploratory questions and hypotheses are then formed concerning both the effects of the device upon musical collaborations, and specific relationships between sensor and music making features. To test these, a confirmatory study is undertaken, which leads to quantitative findings and further research questions. In order to address these questions, a longitudinal, qualitative study is undertaken.

1.3 Thesis Structure

Chapter 2 provides a literature review of background work relating to the measurement of behaviour and affect; and co-present collaborative music making.

Chapter 3 reports an exploratory study, in which various types of sensor data were collected from pairs of improvising percussionists. A multi-level analysis of the data is presented, resulting in a broad set of findings relating to the affective and behavioural measurement of collaborating musicians.

Chapter 4 begins by reflecting upon the results of the first study, and establishing a set of design considerations for a technological intervention. A comprehensive review of related work is then reported. The second half of this chapter describes the design and development of a prototype device – the LuminUs – that provides musicians with real-time feedback on the body motions or glances of their co-performer.

Chapter 5 reports the second study in this thesis, which tested specific exploratory questions and hypotheses concerning i) the effects of the LuminUs device upon collaborating musicians; and ii) relationships between cardiac activity and musical decision making.

Chapter 6 documents the final study in this thesis - a qualitative and longitudinal investigation of the use of the LuminUs by four established musical duos over the course of three rehearsals and a performance.

Chapter 7 draws upon the work undertaken throughout this thesis, and discusses the findings and their implications in relation to existing research.

Chapter 8 concludes this thesis by providing an overview of the work undertaken, and summarising the major findings. It also highlights limitations and areas for future work.

1.4 Contributions

This research contributes to the fields of HCI and affective computing. Both of these fields spawned from the Information Age, and look towards the ways that advances in computing and technology can directly benefit humans. However, the work in these fields has predominantly concerned interactions between humans and computers (e.g. computer gaming), and remote interactions between humans via computers (e.g. mobile communications). This thesis shifts the attention towards *co-present human-human interaction*, and specifically considers the rich interactions that occur between collaborating musicians. In doing so, it evidences ways in which computing and technological advances could enhance these interactions; and consequently demonstrates the importance of considering co-present multi-user scenarios in HCI and affective computing.

Through investigations with collaborating musicians, this thesis also contributes towards human-oriented fields, such as cognitive science, psychology, and music performance. These fields pre-date the Information Age and, whilst researchers have incorporated technological advances into their practice, there is still great scope for exploring the ways in which these advances could assist in the study of humans. In this thesis, sensor technologies are used to collect data from collaborating musicians whilst they perform. Analysis of these data contributes findings relating to the behaviours and experiences of the musicians; and simultaneously demonstrates ways in which sensor technologies can be used as investigative research tools in the aforementioned disciplines.

In addition to these general contributions, specific contributions are detailed below in relation to i) **new methods proposed** for understanding and enhancing affective and behavioural interaction during collaborative music making; and ii) advancing knowledge in multiple disciplines through the **findings obtained**.

1.4.1 New Methods Proposed

Automatic tracking of human-human gaze interactions: Using open source eye tracking glasses and software, along with fiducial markers, a low-cost method is developed and evaluated for automatically detecting gaze interactions between collaborating musicians. To the researcher's knowledge, this is the first time that eye-tracking has been adopted for this purpose. Gaze plays an important role in the non-verbal communication of affective and social signals, and this method provides researchers with a tool for obtaining quantitative measures of gaze interactions.

Extracting quanititative measures of musical decision making: In Chapter 5, two novel methods are reported, applied, and evaluated for the automatic extraction of musical decision making features from MIDI recordings of improvised piano and drum performances. Musical decisions are the decisions that individual musicians make regarding their contributions to an ongoing musical collaboration. Measuring them is especially challenging when the musicians are improvising, due to the fact that musical contributions are not based upon a pre-defined score. The reported methods enable researchers to extract quantitative indicators of the musical decisions made by musicians during improvised performances.

Providing visual feedback of affective and behavioural measures: By designing and testing a prototype device - the LuminUs - this thesis proposes and evaluates a method for providing musicians with real-time inter-personal feedback about the gaze and body motions of their co-performers. To the researcher's knowledge, this is the first time that sensor technologies have been used to provide such feedback to musicians during co-present performances. The reported findings and experiences with the LuminUs serve as an informative reference for researchers and designers involved in the development of devices for providing dynamic visual feedback to performing musicians.

1.4.2 Findings Obtained

Sensing collaborating musicians: Findings reported in Chapter 3 provide empirical evidence for relationships between sensor-derived features and self-reported, or performance-related measures of collaborative music making. In particular:

- Self-reported measures of creativity, engagement, and energy are found to be positively correlated with body motion.
- The number of glances exchanged between musicians are positively correlated with rhythmic synchrony.
- The average glance length is correlated with self-reported boredom.

These findings contribute towards future research by providing an informed indication of which sensors and features might be best suited for the automatic measurement and interpretation of affective and behavioural aspects of collaborative music making.

Musical decision making and cardiac activity: In Chapter 3, relationships between discrete heart rate and musical change events are evidenced. Further findings in Chapter 5 shed light upon relationships between cardiac activity and musical decision making, using both discrete and continuous time series analyses. These results contribute towards the wider literature on psychophysiology and sow the seeds for more in-depth future research on physiological indicators of decision making processes in improvising musicians.

Musicians' use of the LuminUs: Through studies of the musicians' use of the LuminUs, quantitative evidence is obtained for the ways in which sensor-derived, visual feedback devices might have an impact upon musical collaborations. In particular, both gaze and motion feedback are shown to significantly influence the number of glances exchanged between musicians (Chapter 5). In Chapter 6, a longitudinal, qualitative study on the use of the LuminUs identifies and evidences five themes: expression and communication; group attributes; technology; aspects of music; and context. These findings contribute distinct recommendations and considerations for the design and development of affective and behavioural feedback devices for musicians.

Research and technology design recommendations: Based upon the experiences and findings throughout the course of this research, Chapter 7 provides a detailed discussion of recommendations and suggestions relating to the following topics:

- The selection of appropriate sensors for measuring behaviour and affect in collaborating musicians.
- The collection and analysis of dyadic data.
- Implications for the design of new affective and behavioural sensing technologies for collaborative music making.

The discussion of these topics contributes a concise and informative reference for researchers and technologists embarking upon work involving the use of sensors for measuring affective and behavioural aspects of collaborative music making. Additionally, it may also benefit those working within the wider context of the performing arts.

Finally, in an age where the boundaries between art, science, and engineering are becoming increasingly blurred, this thesis contributes to the wider research community by demonstrating the challenges and benefits of undertaking multidisciplinary research.

1.5 Associated Publications

Early motivations and plans for this research are reported in a paper that was presented as part of the Doctoral Consortium at the 2013 conference on Affective Computing and Intelligent Interaction (Morgan et al., 2013). The first study in this thesis (Chapter 3) is reported in a paper presented at the 2014 conference on New Interfaces for Musical Expression (Morgan et al., 2014); and more extensively in an article in the International Journal of Human-Computer studies (Morgan et al., 2015a). The second study (Chapter 5) is partially reported in a short paper, presented at the 2015 INTERACT conference (Morgan et al., 2015b). These publications are listed below:

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. Measuring affect for the study and enhancement of co-present creative collaboration. In *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*, pages 659–664, 2013. ISBN 9780769550480. doi: 10.1109/ACII.2013.115. URL http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6681506

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. Instrumenting the Interaction : Affective and psychophysiological features of live collaborative musical improvisation. In *14th International Conference on New Interfaces for Musical Expression (NIME14)*, London, 2014. URL `http://www.nime.org/proceedings/2014/nime2014_353.pdf`

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. Using affective and behavioural sensors to explore aspects of collaborative music making. *International Journal of Human-Computer Studies*, 82:31–47, 2015a. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2015.05.002. URL `http://www.sciencedirect.com/science/article/pii/S1071581915000853`

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. The LuminUs: Providing musicians with visual feedback on the gaze and body motion of their co-performers. In Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction INTERACT 2015*, volume 9297 of *Lecture Notes in Computer Science*, pages 47–54. Springer International Publishing, 2015b. ISBN 978-3-319-22667-5. doi: 10.1007/978-3-319-22668-2_4

Chapter 2

Background

This thesis takes a multidisciplinary approach, which is influenced by existing research in a number of different fields. This research can be broadly categorised into: i) research relating to the measurement of behaviour and affect, and ii) research dealing with the analysis of co-present, collaborative music making activities. This chapter discusses relevant work within these two categories.

2.1 Measuring Behaviour and Affect

A major component of the research in this thesis is the use of sensor technologies to measure and quantify human behaviour and affect in the context of music making. The tools and techniques that are employed are influenced by research across a range of disciplines, but most prevalently **Psychophysiology**, **Affective Computing**, and **Social Signal Processing**. This section provides a brief introduction to these disciplines, before highlighting literature that is specifically relevant to this thesis.

2.1.1 Relevant Disciplines

Psychophysiology

Psychophysiology is the study of how psychological experiences (thoughts, feelings, emotions) relate to the physiological activity of the body. The typical approach to psychophysiological studies is to measure physiological variables in the laboratory, using equipment developed for medical diagnostics. These measurements are then compared to qualitative and quantitative measures of behaviour and experience, according to the specific focus of the study. Over recent years, equipment for physiological measurement has become increasingly non-invasive, miniaturised and affordable; making it easier to conduct studies, not just in the lab, but also in naturalistic settings (Morgan and Gunes, 2013). Furthermore, these developments are leading towards the integration of physiological sensors in everyday technologies such as phones, smartwatches, and computer game consoles. An advantage of adopting physiological measurement as a means of inferring psychological states is that it often requires no deliberate effort on behalf of the user. The development of a car seat that can sense the driver's heart rate and detect tiredness (Edwards, 2013) is a great example of this paradigm; involving technology development (Walter et al., 2011), research in psychophysiology (Patel et al., 2011), and an application that revolves around the need for non-disruptive measurement.

Affective Computing

The field of affective computing focuses on the development of technologies that are able to recognise, react to, and/or express emotions (Picard, 1997). These emotions are often represented in dimensional space. Commonly used dimensions are valence (unpleasant-pleasant) and arousal (relaxed-aroused) (Gunes and Schuller, 2013). For example, 'excited' would be represented as high arousal and high valence; whilst 'depressed' would be low arousal and low valence. Work in affective computing has mainly focused on categorising the discrete emotional responses of individuals who are presented with pre-recorded, static or virtual stimuli, usually in a laboratory setting (Sariyanidi et al., 2015; Zeng et al., 2009). Researchers are now starting to look towards systems that are able to recognise affect in more true-to-life, spontaneous settings (Gunes and Schuller, 2013). For example, researchers at MIT Media Lab used physiological, behavioural, mobility, and phone usage data to model the happiness of students as they went about their daily lives (Jaques et al., 2015). These data were collected using mobile phones and wrist-worn sensors. In another application, the musical score and sequence of scenes in a film were dictated by the emotional responses of the audience, as inferred from physiological measurements (Price, 2011). These advances are leading to increased commercial interest in affective computing systems; as evidenced by Apple's recent purchase of Emotient (Misener, 2016) - a university spin-off company that specialises in the recognition of emotions from facial expressions.

An area of research that has received less attention is the application of affect recognition in situations where the interaction is occurring *between* people. This, in itself, is not an entirely new idea. In her early work, Picard discussed the potential for affective computing to increase the "affective bandwidth" of person-to-person communication (Picard, 2000, p. 57). However, her choice of language was computer-centric, with the word *bandwidth* implying that emotional communication could be improved by simply transmitting *more* information. More than a decade after Picard made these comments, Janssen et al. (2014) note an absence of work on the automated detection and transmission of emotion during computer mediated, human-human interactions. Furthermore, studies undertaken by Janssen et al. (2014) indicate that the communication of emotion through automated means is perceived as less intimate than user-initiated communication. These findings point towards a need for further research on the underlying nature of effective emotional communication; and the potential applications for affective computing in the context of human-human interactions. In particular, there is a lack of work involving co-present activities; such as collaborative music making, and social interaction.

Social Signal Processing

Social signal processing (SSP) (Pentland, 2005) is a comparatively new domain, which aims to provide computers with the ability to sense and understand human social signals (Vinciarelli et al., 2009). In this context a social signal is defined as a "communicative or informative signal that, either directly or indirectly, provides information about 'social facts', that is, about social interactions, social emotions, social attitudes, or social relations" (Pantic et al., 2011, p. 9). SSP is relevant to the work in this thesis because it is inherently concerned with phenomena that occur during human-human interactions. For example, Curhan and Pentland (2007) investigated how conversational dynamics, measured by extracting speech features, could be used to predict the outcomes dyadic negotiations. With respect to emotion, SSP takes a markedly different approach to affective computing research. The latter has predominately investigated basic emotions that can be experienced individually, such as happiness and anger; whereas SSP considers social emotions, such as empathy and envy (Vinciarelli et al., 2012). Furthermore, the analysis of complex *non-verbal* behavioural cues, such as gesture and gaze, is a central focus of SSP. This makes it particularly relevant to the study of collaborative music making, where the lack of verbal communication places emphasis upon non-verbal modalities.

A common theme within these disciplines is that human experience and affect can be segmented into three components; cognitive (thoughts), behavioural (expressions and actions) and physiological (biochemical and electrical changes in the body). Measuring each of these components presents varying challenges, requiring distinct technologies and processing techniques. The following sections discuss relevant research relating to each component, since the studies in this thesis incorporate measurements of all three.

2.1.2 Cognitive Measurement

Information about cognitive processes inside the brain can be obtained using neuroimaging techniques. The most commonly used techniques are electroencephalography (EEG) and functional Magnetic Resonance Imaging (fMRI). Both techniques have been used to identify emotions by analysing the brain's response to affective stimuli such as images (Schneider et al., 1997) and music (Schmidt and Trainor, 2001). fMRI has a higher spatial resolution than EEG, but the requirement for participants to be stationary inside a large and immobile scanner makes it impossible to conduct studies in naturalistic settings. For example, in a study of the neural aspects of musical improvisation, jazz pianists were asked to play whilst lying in an fMRI scanner (Limb and Braun, 2008).

EEG data are collected by placing electrodes on the subject's scalp. Affordable and user-friendly EEG systems, such as the Emotiv EPOC[1] headset, have become available in recent years. The Emotiv system has been used in the classification of affective states (Ramirez and Vamvakousis, 2012) and the detection of mental actions (Taylor and Schmidt, 2012). However, it has been shown to under-perform in comparison to medical-grade devices (Duvinage et al., 2013).

A major issue with all current EEG devices is that the recordings are highly susceptible to movement-related artifacts. This poses difficulties for data collection from performing musicians (Maidhof et al., 2014). Despite this issue, Babiloni et al. (2012) used EEG sensors to investigate empathy in musical quartets. They found relationships between empathy scores and frontal oscillatory alpha rhythms collected while the musicians observed a recording of their performance. No equivalent relationships were found for EEG data collected during the performances. Sänger et al. (2012) collected EEG recordings from guitar duets and found evidence to support the claim that synchronous oscillatory brain activity plays a functional role in musical performance. They also found that phase locking (a measure of synchronisation) was modulated in relation to the leader and follower roles assigned to the musicians. Both of these studies involved musicians performing prepared musical pieces. To the best of the researcher's knowledge, there have been no studies involving the collection of cognitive measures during collaborative musical improvisation.

[1] `https://emotiv.com/`

2.1.3 Behavioural Measurement

The research in this thesis concerns the measurement of relatively short-term behaviours (in the order of seconds and minutes), many of which can be categorised as non-verbal communicative acts. Argyle (1978) outlines seven forms of non-verbal communication (NVC): *facial expressions, gaze, gestures, bodily posture, bodily contact, spatial behaviour* (e.g. proximity), and *appearance* (e.g. clothing). He models NVC as a simple, communications theory-inspired sequence, whereby a sender encodes a 'social signal', which is subsequently decoded by a receiver. It is implicit that this signal transfer is not error free. One person will never perfectly interpret the non-verbal communicative act of another person. Furthermore, the process of NVC does not always involve conscious awareness. This makes it a particularly interesting parameter in the study of human interactions, as it indicates subtle features of the interaction that may not be revealed by self-report measures.

In the context of co-present musical interactions, gaze, bodily posture, and spatial behaviours are of particular interest. Gaze can reveal information about the dynamics and nature of co-present human interactions. Numerous studies have shown how gaze is closely synchronised with speech during conversations (Kendon, 1967; Cummins, 2012; Oertel et al., 2012). Additionally, the amount of time people spend looking at each other has been shown to relate to dominance and rapport (Argyle, 1978). Mutual gaze has also been shown to be physiologically arousing (Mazur et al., 1980). Gaze is commonly measured by manually annotating video footage, which is a labour-intensive task (Ye et al., 2012). However, modern eye-tracking glasses are able to continuously track where someone is looking within a scene captured from a head-mounted camera (Kassner et al., 2014).

With respect to bodily posture and spatial behaviour, accurate measurements can be obtained using optical marker-based motion tracking systems, which use multiple cameras to detect small reflective markers positioned on the body. Glowinski et al. (2013) used such a system to study the bodily movements of a string quartet. Their results suggest that head movement features can be used to distinguish between an engaging and non-engaging performance (as rated by the performers). Healey et al. (2005) used marker-based tracking to examine the spatial behaviour of a group of seven improvising musicians. They observed how the use of space played a complex role in maintaining the coherence of the performance, and drew a number of parallels with conversational interactions. The trade-off with marker-based systems is that they take some time to set up and are not particularly portable (Bianchi-Berthouze and Kleinsmith, 2015). With the advent of the Microsoft Kinect, researchers have been able to measure body movements using a single depth-sensitive camera (Hadjakos et al., 2013; Morgan and Gunes, 2013). This option offers a low cost, portable, and unobtrusive option for tracking motion in naturalistic settings. However, the tracking accuracy can be poor (Dael et al., 2016). An additional option is to record kinematic features of body motion using handheld or wearable accelerometers. If multiple accelerometers are used then the data can be processed using kinematic models in order to estimate spatial positioning (Zhou et al., 2008)

It should be noted that the non-verbal behaviours defined by Argyle, and discussed above, are all forms of *visual* communication. In the study of musical interactions it is also important to consider the effects that auditory feedback might have on NVC between musicians. For example, during collaborative improvisation musicians frequently communicate using auditory cues, such as playing louder (Jensen and Marchetti, 2010). Nowicki et al. (2013) found that the timing synchrony between two musicians tapping a rhythm improved when the taps triggered sounds. Given that the sounds a musician creates are directly related to their playing, auditory-feedback related measures can be collected using Musical Instrument Digital Interface (MIDI)-enabled instruments. MIDI data provides accurate information about the timing and velocity (loudness) of musical notes. Coupled with the other measures discussed in this section, a wide range of options exist for collecting rich data relating to the behaviours of interacting musicians.

2.1.4 Physiological Measurement

Physiological measures can be used to capture subconscious and spontaneous aspects of a person's cognitive and affective states (Fairclough, 2009; Cacioppo et al., 2007b). These measures can be collected whilst a person goes about their daily activities (Hernandez et al., 2014), which makes them particularly suitable for the study of collaborating musicians. Physiological data are commonly collected by measuring mechanical, electrophysical and biochemical signals, using surface electrodes and sensors positioned at specific sites on the body (Cacioppo et al., 2007a). In a study of flow[2] during piano playing, de Manzano et al. (2010) measured heart rate, respiration and facial muscle movements while professional pianists gave five performances of a prepared piece. For each performance the pianists were subsequently asked to rate their level of flow using a questionnaire. A significant relationship was found between flow and heart

[2]Csikszentmihalyi's theory of Flow (Csikszentmihalyi and Rathunde, 1993) describes the mental state of being completely absorbed in an activity.

rate variability, respiratory depth, and facial muscle movements. Physiological studies have also been conducted to try to measure musicians' emotions during musical performance (Knapp et al., 2009). A potential application of live physiological measurement in music collaboration has been demonstrated by Mealla et al. (2011), who created an interactive musical tabletop, where physiological signals contributed to the generated sounds.

In a study of the physiological reactions of audience members to a live music performance, Egermann et al. (2013) found that unexpected musical events were generally associated with a rise in skin conductance, and decreased heart rate. A real-world application of these ideas saw the musical score and sequence of scenes in a film being dictated by the emotional responses of the audience, as inferred from physiological measurements (Price, 2011).

Regarding human-human interactions, a study of partner influence during conversation found 'physiological linkage' between the blood pressure (BP) measurements of romantic couples (Reed et al., 2013). Research into user experience with game technologies found differing physiological responses when participants were playing against a computer compared with playing against another human (Mandryk et al., 2006).

In summary, cognitive, behavioural and physiological measurements provide many ways to infer and quantify affective aspects of human experience. In recent years the use of these techniques has started to make the transition from laboratory-based studies to real-world applications. However, there has been comparatively little focus on situations where the interaction is occurring *between* people.

2.2 Co-Present Collaborative Music Making

In the context of this thesis, co-present collaborative music making is defined as any situation where two or more people are jointly involved in the process of creating live music. The following sections discuss research that is relevant to **group interactions** in general; and **group musical improvisation**, which is a specific case of co-present collaborative music making that is investigated in this thesis. The final section provides a review of approaches to **studying co-present musical interactions**.

2.2.1 Group Interactions

Sawyer (2003) proposes two generalised approaches to the study of group interactions: the *input-output approach*; and the *process approach*. The former concerns things that take place before and after the interaction, whilst the latter looks at what occurs during the interaction. Quantitative methods can provide interesting insights into group interaction processes. For example, Fencott and Bryan-Kinns (2010) used interaction logs to analyse specific aspects of software-based co-present collaborative music making, such as the amount of co-editing that occurred. Hadjakos et al. (2013) developed a quantitative method for analysing the rhythmic synchronisation of a violin duo using the Kinect motion tracking device. By tracking head movements they were able to demonstrate how complex interaction patterns could be observed.

As an extension to the theory of flow (see section 2.1.4), Sawyer (2003) conceives the idea of group flow, referring to a state of peak performance at the level of the group, rather than the individual. He points to the importance of factors such as parallel processing (simultaneous awareness of self and collaborator(s)) and visual attention in establishing a state of group flow. However, evidence for the theory of group flow is somewhat anecdotal. Indeed, Sawyer suggests that the study of group flow "must proceed by examining the interactional dynamics among members during performance" (Sawyer, 2003, p. 47). A similar concept is that of mutual engagement, which has been highlighted as an important feature of group musical interactions (Bryan-Kinns, 2013). Bryan-Kinns and Hamilton (2012) developed a mutual engagement questionnaire for evaluating the mutual engagement qualities of different musical interfaces.

Emotion also plays an important part in group interactions. Of particular interest are the processes by which the emotional representations of a person or group influence the emotions of another person or group. This is often termed *emotional contagion*. There is a lack of research on emotional contagion in performing musicians; however, it has been suggested that emotional contagion is one of the mechanisms by which people experience emotions while listening to music (Lundqvist et al., 2008). Studies of group interactions have shown that emotional contagion affects group processes, such as task performance and cooperation (Barsade, 2002). Similar processes of influence have been observed in relation to behavioural displays. *Behavioural mimicry* is the process whereby the actions or emotions represented by one person subconsciously cause congruent behaviour in another person. There is a growing body of evidence for behavioural mimicry and its links to cognitive processes (Chartrand and Lakin, 2012). In the musical domain it has been suggested that behavioural mimicry contributes to temporal assimilation and coordinated variations in intensity and intonation during ensemble performances (Keller et al., 2014). By collecting and analysing sensor measurements from collaborating musicians, quantitative evidence could be obtained to investigate the theories of group flow, emotional contagion, and behavioural mimicry, discussed in this section.

2.2.2 Group Musical Improvisation

Group musical improvisation is defined as a spontaneous process, whereby creative contributions are made within the restrictions of real-time performance (Kenny and Gellrich, 2002; Wilson and MacDonald, 2012). Research on group musical improvisation is limited, and existing studies have predominantly focused on jazz music (Wilson and MacDonald, 2012). Seddon (2005) used videotapes of six jazz musicians during rehearsal and performance in order to investigate modes of communication during jazz improvisation. He defined three modes of non-verbal communication: *instruction*, *co-operation*, and *collaboration*. Instruction involves the demonstration of musical ideas through vocalisation, or use of an instrument. Cooperation occurs when the musicians are producing a cohesive performance with the inclusion of musical and visual cues. Collaboration is the state where musicians are able to stimulate the spontaneous generation of creative contributions within the group; communicating exclusively through musical interaction. During collaboration, Seddon defines the musicians as being 'empathetically attuned' - a state of mutual empathy that encourages the musicians to take risks and challenge each other's creativity.

Regarding the cognitive factors involved in improvisation, Kenny and Gellrich (2002) propose eight processes: short, medium, and long-term *anticipation*; short, medium, and long-term *recall*; *flow status* (see section 2.1.4); and *feedback* (decisions based upon previous experiences). Similar cognitive processes were also identified by Biasutti and Frezza (2009), who developed an improvisation process questionnaire, which they gave to 76 experienced musicians. In addition to anticipation, use of repertoire, flow, and feedback, the authors highlight *emotive communication* as an important ability required by improvisers.

Comparisons have been made between jazz improvisation and conversation (Monson, 1996; Sawyer, 2005), whereby a common language or vocabulary is used alongside rules (e.g. grammar) in order to form a coherent and emergent interaction between two or more people. Conversation analysts have described how interlocutors use non-verbal behaviours such as eye gaze (Kendon, 1967), and body position (Kendon, 2010) to maintain successful conversations. There is already some evidence that similar processes may occur during collaborative musical creativity (Healey et al., 2005). A noteworthy difference between conversation and musical improvisation is that the former is characterised by turn taking (Sacks et al., 1974), whilst the latter involves simultaneous contributions. Musicians must, therefore, continuously monitor the contributions

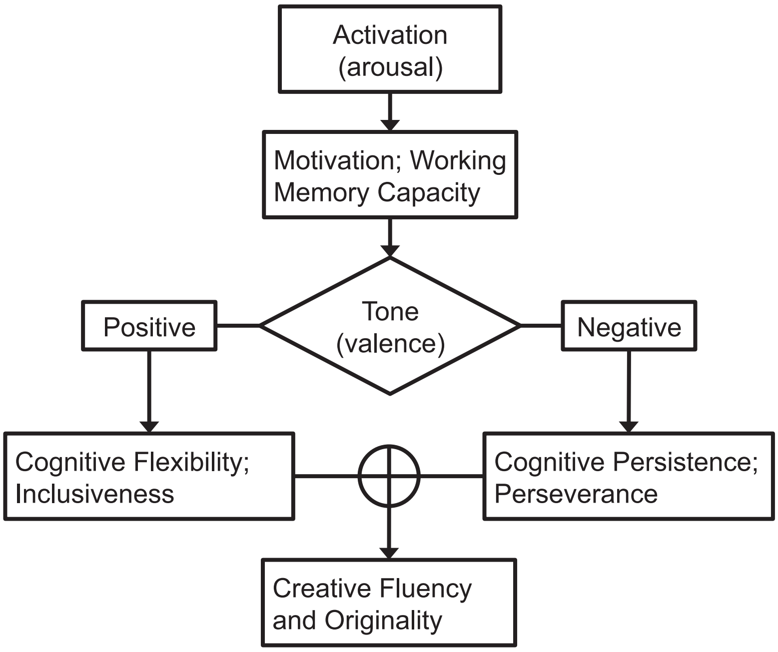


Figure 2.1: Schematic of the dual pathway to creativity model (reproduced from De Dreu et al. (2008)).

of their collaborators, whilst also providing their own novel contributions. This involves a combination of both conscious and sub-conscious processing (Sawyer, 2003).

Finally, it is worth considering how the moods and emotions of group members might influence the creative aspects of improvisation. A theory for the relationship between emotion and creativity is provided by the 'dual pathway' model (De Dreu et al., 2008) (see Fig. 2.1). This model suggests that emotions with high arousal (e.g. anger, elation) lead to greater originality and creative fluency (the number of ideas, insights and solutions generated) when compared to low arousal emotions (e.g. sadness, serenity). It also proposes that positive emotions contribute to this process by facilitating greater cognitive flexibility and inclusiveness; whilst negative emotions facilitate increased persistence and perseverance. It is important to note that this model was developed based upon studies of individual creativity, and it is not clear how it might apply to group creativity. Existing studies of group creativity have predominantly explored the influence of long-term mood on creative output, often in a workplace setting (Amabile et al., 2005; Jamison, 1996). A notable feature of these studies is that they tend to focus on correlations between discrete emotional states and overall creativity. There is an absence of research that addresses continuous affective interactions and their real-time influences on creative tasks.

2.2.3 Studying Co-present Musical Interactions

Quantitative studies of co-present musical interactions commonly adopt repeated measures designs; where musicians interact under contrasting experimental conditions. These studies tend to involve short-term musical interactions, which are performed in controlled settings. For example, Vera et al. (2013) used three visibility conditions – full, partial, and no visibility – to investigate the effects of co-performer visibility upon synchronisation during ensemble violin performances. A similar study design was used by Goebl and Palmer (2009) to investigate the influence that auditory feedback had upon the synchronisation of note timing and head motion between pairs of pianists. In an investigation of emotional entrainment, Varni et al. (2008) measured the head movements of violinists whilst they played duets in different emotional state conditions.

Longer-term studies often take a more qualitative approach, involving the use of observations and interviews. For example, Schiavio and Hoffding (2015) report a qualitative account of joint musical performance, based upon semi-structured interviews conducted with members of the Danish String Quartet after two musical tours. Biasutti (2012) investigated group music composing strategies by analysing video recordings of a rock band composing a new piece over seven sessions.

In HCI-related fields, a common motivation for musical interaction studies is the investigation of new technologies and interfaces. Gloor et al. (2013) used 'sociometric badges' containing accelerometers and microphones to measure the energy levels of a jazz band whilst they performed in front of an audience. This enabled them to investigate group flow, by looking at the synchronisation of energy levels between band members. Bengler and Bryan-Kinns (2013) developed an interface for collaborative musical performance and evaluated it in a public setting using questionnaires, observations, and data logged from the interface. In this case, the data also served to inform wider implications for the facilitation of collaborative music making. Stowell et al. (2009) review various approaches to the evaluation of live music-making interfaces and highlight the importance of designing experiments that reflect an authentic interaction context.

The studies discussed in this section supplement those reviewed throughout this chapter; emphasising the range of different approaches to studying co-present musical interaction. When designing a study, the researcher must consider multiple factors, such as then length of the study; the type of data to be collected; and whether to use controlled or naturalistic settings. These considerations are made in view of the aims of the research and the resources available.

2.3 Summary

This chapter reviewed existing work relating to the measurement of affective and behavioural features of collaborative music making. This work was categorised according to two topics: i) measuring behaviour and affect; and ii) co-present collaborative music making. Despite this distinction, there are three underlying themes, which are central to the research conducted in this thesis: *technologies*, *methods*, and *concepts*. We have seen how motion sensing technologies can be used for tracking the subtle movements of musicians and measuring engagement; whilst small, wearable physiological sensors are used to sense and recognise affective states and social signals. A range of methods for the collection and analysis of data have been described. These include the analysis of alpha rhythms extracted from the EEG data of musical ensembles; the analysis of heart rate variability in relation to the flow experiences of pianists; and the use of interviews and observations for investigating long-term musical interactions. Finally, relevant concepts have been discussed. These include the theory of group flow; the dual pathway approach to creativity; and concepts surrounding non-verbal communication.

In the following chapter an exploratory study is documented, which draws upon these technologies, methods, and concepts. In particular, physiological, motion and MIDI sensors were used to collect data from improvising percussionists during controlled experiments. This rich data-set is subsequently analysed and discussed in relation to existing concepts, such as flow and the dual pathway model.

Chapter 3

Study 1

Sensing Collaborating Musicians

This chapter documents an exploratory study, which was designed to investigate the use of various sensor technologies for measuring affective and behavioural aspects of collaborative music making. Exploratory studies provide a means of revealing previously unimagined connections and causal mechanisms; and are well suited to cases where little is known about the area of interest (Reiter, 2013). Informed by the methods and techniques discussed in the previous chapter, subjective and continuous quantitative measures were collected from pairs of co-present, improvising drummers. To investigate the importance of visual contact between participants, two conditions were used: one where the participants were visible to each other; and one where they were separated by a screen. A broad analyses of the data are undertaken, and findings are reported at the level of the individual, dyad, and entire study group.

3.1 Aims and Motivation

This study was not founded on specific hypotheses, instead the aims were to:

Assess the challenges and issues associated with the experimental use of behavioural and affective sensors to investigate collaborative music making.

Report exploratory findings to guide and inform future research.

Identify which measures and features are most suitable for the investigation of collaborative music making.

The study was inspired by the research described in the previous chapter, and by the broader motivations of this thesis, set out in Section 1.1. The decision to investigate dyads was made because dyadic grouping is the simplest instance of collaborative interaction, which avoids some of the practical challenges and complexities that could be associated with investigating larger groups, such as quartets. Furthermore, this provides a basis upon which to consider investigations of larger groups.

In contrast to playing pre-composed music, co-present improvisation was chosen because it presents particular challenges to the musicians, requiring them to pay attention to each other's ongoing contributions and communicate non-verbally in order to establish a successful interaction.

The selection of measures to be tracked was motivated by four factors: i) the availability of sensor technologies; ii) the practicalities of collecting the measures in a live performance scenario; iii) the measures used in existing studies (reviewed in the previous chapter); and iv) the researcher's own previous experience of collecting physiological measures from performing musicians during his Master's research project (unpublished).

Few assumptions were made about which non-verbal aspects of co-present musical collaboration might contribute towards the success of the performance and engagement of the participants. However, the decision to use visual and non-visual conditions was motivated by a desire to investigate the influence of non-verbal communication conveyed through the visual channel. Previous studies have highlighted the importance of visually conveyed non-verbal signals in musical collaboration (Davidson and Good, 2002; Vera et al., 2013).

Drumming was chosen in this study because it presents some noteworthy advantages over other forms of musical expression. In particular, beat timing and velocity can be accurately recorded using electronic pads; and large amounts of motion are involved, which increases the information conveyed visually, through movement. There is also an absence of melodic content, which might otherwise influence the participants' affective responses. The experiment was simplified further by requiring that each participant only used one hand to drum on a single drum pad. This also meant that the non-drumming hand could be used to collect physiological data, using sensors attached to the fingers. The following sections provide more detailed descriptions of the study design and data collection methods, as well as the steps involved in processing the data.

3.2 Research Design and Data Collection

A within-subjects design was used, where each pair of drummers performed drumming tasks under each of the two conditions: visual (V) and non-visual (NV). This design was selected due to high individual variability in drumming ability, behaviour, and affective responses. Throughout the remainder of this chapter the word 'dyad' is used to refer to a specific pairing of two drummers. Each drummer within a dyad is distinguished as participant 1 or 2. The notation Dx.py.C is used to indicate the dyad (D), participant (p), and condition (C).

3.2.1 Participants

Participants were recruited via email lists and word of mouth. All participants were required to have prior drumming experience, to the extent that they were confident enough to improvise rhythms. Five pairs of participants took part in the study (2 mixed-sex pairs, 3 male pairs). Participants in each dyad knew each other, and three of the dyads had previously played music together. The choice to pair people who knew each other was specifically made in order to maximise the amount of non-verbal communication that could be observed. However, it is beyond the scope of this study to include inter-participant relationship variables as part of the analyses. The participants were aged 26 to 34 (M=29.1, SD=3.1), they had been playing percussion for between 1 and 17 years (M=7.4, SD=5.0), and their self-rated level of expertise ranged from 2 to 4 (M=2.7, SD=0.7) on a five point scale representing novice (1) to expert (5). Participants were offered £20 as an incentive and signed a consent form before partaking in the study. The study was given ethical approval (QMREC2013/48)[1].

3.2.2 Measures

The choice was made to collect a wide range of measurements so that different techniques could be evaluated and correlations between various types of data and extracted features could be explored. Small wireless electrocardiogram (ECG) and galvanic skin response (GSR) sensors (53 mm × 32 mm × 19 mm) were used to measure heart rate and perspiration respectively. These were developed by Shimmer[2]. The Emotiv EPOC EEG headset was used to wirelessly record 14 channel EEG measurements from each participant. All the physiological sensors contained accelerometers for recording motion. For the drums, two identical Roland V-Drum electronic drum pads were used. By recording MIDI data from the pads it was possible to log the exact timing and velocity (loudness) of each drum beat. Three video cameras were set up: one facing each participant, and one overhead camera to capture the entire interaction. Sample

[1] Queen Mary Research Ethics Committee reference number.

[2] `http://www.shimmersensing.com/`

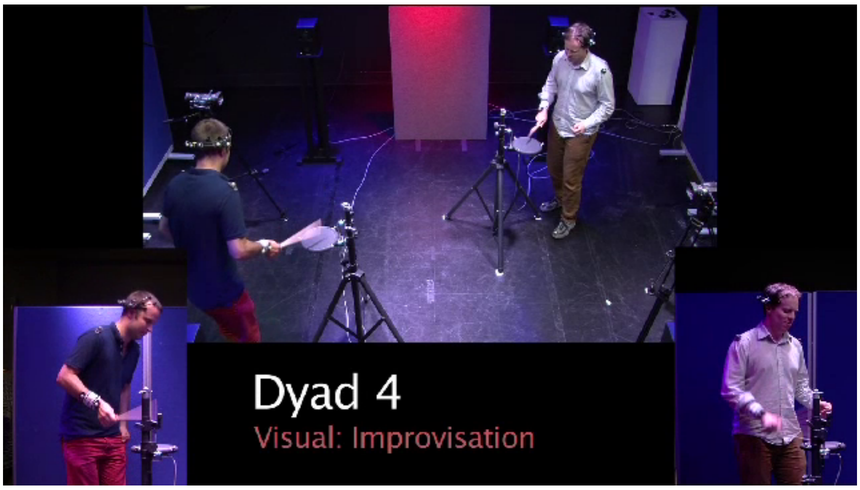


Figure 3.1: A still image of the captured video. Two minute video excerpts are available online at `http://dx.doi.org/10.1016/j.ijhcs.2015.05.002`.

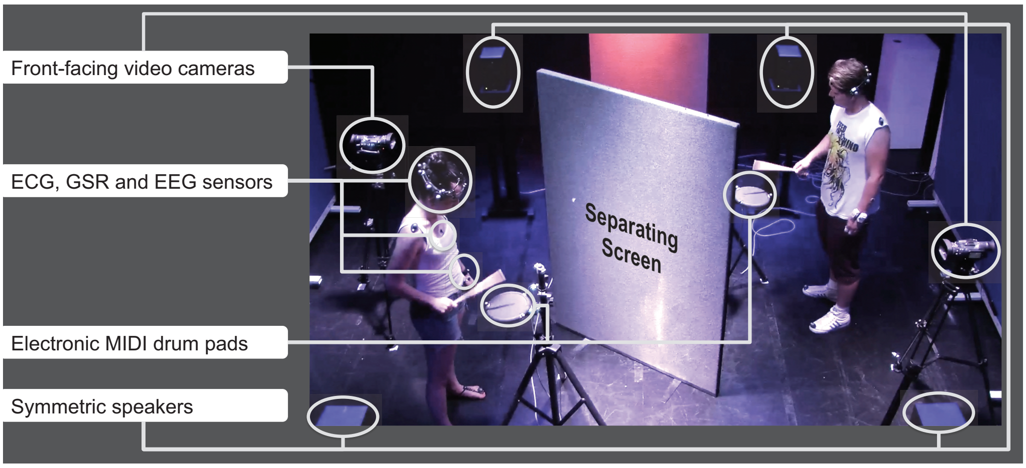


Figure 3.2: Image taken from the overhead camera illustrating the setup and the equipment used in the study.

video footage of the sessions can be obtained online (see Fig. 3.1). Figure 3.2 shows an annotated image taken from the overhead camera.

A post-performance questionnaire (PPQ) was designed to collect self-report (SR) data from each participant while they reviewed video footage of their improvised performances. The PPQ asked participants to rate their individual levels of creativity, engagement with the other participant, energy, positivity and boredom on a 9-point scale; as well as who they thought was leading the performance (1 = 'All me', 9 = 'All them'). The full PPQ is shown in Fig. 3.3. The first item is loosely based upon the widely used consensual assessment technique (CAT) (Amabile, 1982), which proposes that the best way to measure the creativity of an artefact is to simply ask experts in that field to provide a creativity rating. With respect to this study, there was a specific interest in measuring the participant's own subjective ratings of their creativity. Previous studies have found moderate correlations between self and expert-rated creativity using the CAT (Hennessey et al., 2011). Item 2 is adapted from the mutual engagement questionnaire (Bryan-Kinns and Hamilton, 2012), developed specifically for musical interactions. Items 3, 4 and 6 are closely related to the items activation, valence, and dominance from the self assessment manikin (SAM), which is commonly used in affect research (Gunes and Schuller, 2013). Item 5 is adapted from a questionnaire used to study user experience with entertainment technologies (Mandryk et al., 2006; Mandryk and Atkins, 2007).

How creative/original was the drumming?	Not at all	Very
How engaged were you with the other participant?	Not at all	Very
How energetic did you feel?	Not at all	Very
How positive did you feel?	Not at all	Very
How bored did you feel?	Not at all	Very
Was someone leading the performance?	All me	All them

Figure 3.3: The post-performance questionnaire (PPQ) that was used in the study.

3.2.3 Data Synchronisation

Two computers and three separate applications (MATLAB, Emotiv TestBench, Logic Pro) were used to record the continuous measurements. To synchronise these data, physiological and EEG sensors were placed on top of one of the drum pads, which was subsequently hit 10 times. This meant that there were 10 clearly identifiable, short-duration peak events in the EEG and physiological accelerometer data, accompanied by 10 MIDI note events and 10 visible video events. When processing the data, these events were used as reference points, allowing all the data sources to be aligned to high (millisecond) precision.

3.2.4 Setup

The study was held in a room designed for performance studies, with stage lighting set up to make it feel like a live music venue. The drum pads were positioned in the centre of the room, with speakers either side (see Fig. 3.2). The two computers were placed behind blank screens at one end of the room; this was also where the experimenter sat during the performances. ECG modules were strapped around each participant's waist, with the electrodes attached to their chest. GSR modules were placed around the wrist of their non-drumming hand, and the electrodes were strapped to their index and middle finger. The EEG headsets were placed on the participants' heads, and electrodes were individually adjusted to obtain a good signal (as indicated by the software).

3.2.5 Tasks

The experiment consisted of two warm up tasks followed by an improvisation task. These three tasks were performed twice, once under a non-visual (NV) condition, then again under a visual (V) condition. In the NV condition the participants faced a blank screen so that they were unable to see each other. In the V condition they faced towards each other with no obstruction, other than the drum pad. The first warm up task ($\sim$ 1 min) required the participants to hit their drum in synchrony with a metronome click track at a tempo of 110 beats-per-minute (bpm). The second warm up task ($\sim$ 1 min) required them to repeat a set rhythmic phrase, which they listened to and learnt prior to the task. These initial tasks allowed the participants to get used to playing the electronic drum and to drumming with one another. The improvisation task (*Improv*, $\sim$ 6-10 min) required the participants to improvise with one another, where the only restriction was that they did not use verbal communication. Verbal communication was prohibited due to a desire to simulate a live performance environment, where musicians would not normally use verbal communication. Following completion of the drum tasks, the participants sat individually and watched the overhead videos of their two improvised performances. After each minute[3] of video they were asked to complete all the SR items on the PPQ, in relation to that particular minute of their performance.

3.3 Data Processing

3.3.1 Preparation

The EEG, ECG, GSR, MIDI, and self-report (SR) data were imported into MATLAB. For each dyad the accelerometer synchronisation peaks and MIDI note events were used to align the data to a common start point (t_0). Using the video footage, the start and end times of each task were found, relative to t_0. For each data source these time

[3]For dyad 5, 2 minute segments were used because the *Improv* tasks were longer.

points were used to extract and label data-sets corresponding to measurements for each participant within each task and visibility condition.

3.3.2 Feature Extraction

Features were extracted from individual data-sets according to the type of data they contained. The data from the *Improv* tasks was initially segmented according to the 1 or 2 minute time windows used for the SR questionnaires. Features were then extracted from each of these windows ($Improv_w$), for each participant within each condition (NV or V). This was done to facilitate the analysis of relationships between extracted features and SR measures, as detailed in the following section. Some of the EEG, ECG, and GSR data were found to contain artifactual readings, due to movement and poor electrode contact. These noisy data were manually labelled using a binary coding (1 = noisy, 0 = clean), so that they could be recognised and automatically excluded from the subsequent analysis.

ECG: ECGtools[4] was used to filter the raw ECG data (sampled at 51.2 Hz) and extract the R-peaks, which correspond to individual heart beats. The distance between consecutive peaks was then used to find the instantaneous heart rate (HR) values. These values were interpolated to give an evenly spaced time series, from which the following features were extracted: mean, SD, maximum, minimum, the positions of maxima and minima, and the number of extrema divided by the task duration.

GSR: Skin conductance response (SCR) has been shown to be a useful metric in analysis of GSR data (Kim et al., 2004; Kim and André, 2008). Ledalab[5] was used to extract the timing and amplitude of SCR events from the raw GSR data (sampled at 5 Hz) using continuous decomposition analysis (CDA). Again, interpolation was performed and the mean, SD, positions of maxima and minima, and number of extrema divided by task duration, were calculated from the SCR amplitude series.

EEG: EEG signals contain frequency components that relate to the firing activity of neurones in the brain. Standard Theta, Alpha, and Beta frequency band power values are often computed in EEG studies, as they provide information on cognitive activity (Chaouachi and Frasson, 2010). Using EEGlab (Delorme and Makeig, 2004) the EEG signal (sampled at 128 Hz) was initially bandpass filtered between 3 and 30

[4] `http://www.ecgtools.org/`
[5] `http://www.ledalab.de/`

Hz. Manual artifact rejection was then performed to remove noisy segments of data caused by head and facial muscle movements. The Emotiv EEG recordings consist of 14 channels of data, relating to sensors at different positions on the scalp. Artifactual channels were removed entirely, and the average power over all remaining channels was computed within the following standard frequency bands: Theta (4-7 Hz), Alpha (7.5-12.5 Hz), L-Beta (12.5-25 Hz), and H-Beta (25-30 Hz).

Motion: The accelerometer readings were taken from the ECG, GSR and EEG sensors and the absolute values of the axial components were summed for each sensor. This provided approximate mean quantity of motion (QoM) values for the head (EEG), torso (ECG), and non-drumming hand (GSR). QoM has previously been shown to be one the most successful motion features for classifying gestural representations of emotions (Castellano et al., 2007).

MIDI: The number of beats per second, SD in time between consecutive beats, and mean velocity were computed as MIDI features. The beat onset times were compared for participants in each dyad ($tP1$ and $tP2$), and any beats that occurred within 70 ms of each other were considered to be perceptually synchronous rhythmic events, as suggested by Dixon (2001). For these beats, the time difference was calculated ($tP1$ - $tP2$). The mean over all the absolute difference values was then calculated in order to provide an indicator of the timing synchrony within the dyads. The MIDI data were also used to manually annotate the rhythmic change points (RCPs) for each participant. RCPs were defined as the time points in seconds (relative to t_0) when the participant changed their rhythm from a previously established pattern: defined as a fixed-length sequence of beats, which had been repeated at least twice. The identification of RCPs is illustrated in Fig. 3.4, which shows a section of MIDI data from one of the participants. As can be seen, the position of the MIDI notes on the score, as well as the velocity of the notes (denoted by colour) helps to identify two distinct rhythmic patterns. In addition to visual analysis, aural analysis of the rhythms also facilitated the identification of RCPs.

Video: The use of video footage as an elicitation tool to help the participants provide PPQ responses has already been described in Section 3.2.2. In existing studies of conversational interactions video footage has also been used to annotate gaze behaviour (Oertel et al., 2012; Cummins, 2012). In this study the footage from the front-facing video cameras was used to manually annotate the time points when a par-

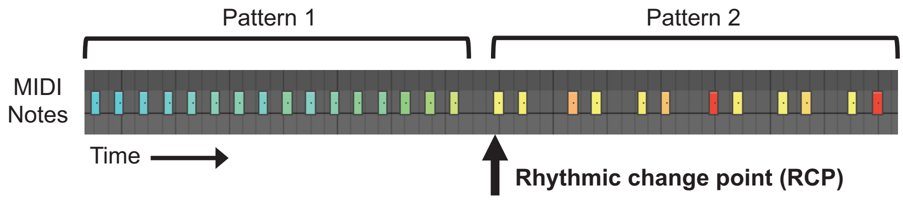


Figure 3.4: A short segment of MIDI data for one of the participants, illustrating how rhythmic change points were identified. The blocks represent individual beats, and their colour denotes the velocity (strength) of the beat. There are two distinct patterns. Pattern 1 consists of soft (blue-green) and evenly spaced beats, and Pattern 2 consists of paired beats, played with more velocity (yellow-red).

ticipant was glancing towards their collaborator during the *Improv* tasks. The start and end time of each glance was recorded in seconds (relative to t_0). These data were then used to calculate the number of glances, percentage glance time, number of mutual glances, and percentage mutual glance time, within each time window ($Improv_w$).

Table 3.1: A summary of the measures, sensors, and features used in this study.

Measure	**Sensor**	**Features**
Self Report	PPQ responses (video elicited)	Creativity, engagement, energy, positivity, boredom, leadership. (see Fig. 3.3)
ECG	Shimmer ECG sensors	Heart rate (HR); mean and SD of HR; positions of HR extrema; number of extrema per min.
GSR	Shimmer GSR sensors	Skin conductance response (SCR); mean and SD of SCR; positions of SCR extrema; number of extrema per min.
EEG	Emotiv EEG headset	Mean power in Theta (4-7 Hz), Alpha (7.5-12.5 Hz), L-Beta (12.5-25 Hz) and H-Beta (25-30 Hz) bands.
Motion	Accelerometers in the Shimmer sensors	Mean quantity of motion (QoM - head, torso, hand).
MIDI	Roland V-Drums	Beats per second; SD in time between consecutive beats; mean velocity; mean absolute inter-beat lag between participants; rhythmic change points (RCPs).
Video	Front-facing and overhead cameras	Number of glances; % glance time; number of mutual glances; % mutual glance time.

3.4 Analyses and Results

The aim of the analyses was to use the rich data-set to undertake an exploratory investigation of dyadic, collaborative music making. To facilitate this the analyses were structured over three levels: i) **individual**; ii) **dyad**; and iii) **whole study**. This allowed the analysis to focus upon general trends and observations over the entire study whilst also paying attention to interesting and suggestive idiosyncrasies at the level of the individual and the dyad.

3.4.1 Individual-Level Analyses

Data Visualisation

Of particular interest at the level of the individual was the way in which specific events within the improvisation tasks might be linked to observable changes in the continuously captured physiological and motion-based measures. This interest was inspired by previous studies identifying patterns of physiological change in response to musical events (Egermann et al., 2013), and computer game events (Ravaja et al., 2006). Within the context of the present study, the two main event based measures that were obtained were the *rhythmic change points (RCPs)*, and *phases of glance* (see Section 3.4.3) for each participant. These are referred to as continuous features, since they derive from continuous measures: audio, midi, and video. Figure 3.5 illustrates a simple time series plot that was designed to enable the visualisation of these events alongside other continuous features. By generating such a plot for each dyad within each condition, it was possible to visually explore potential patterns within, and relationships between, continuous measures (plots for all dyads can be found in Appendix A.3).

Relationships Between Heart Rate and RCPs

One of the most interesting observations to come from the visual analysis was that RCPs for three of the participants seemed to be closely aligned with extrema (peaks and/or troughs) in the participant's heart rate plot. To test whether this relationship was significant, heart rate extrema time points were extracted for each participant, within each condition, and the corresponding number of aligned RCPs was calculated. Points were arbitrarily defined as *aligned* if they fell within a quarter of a second of

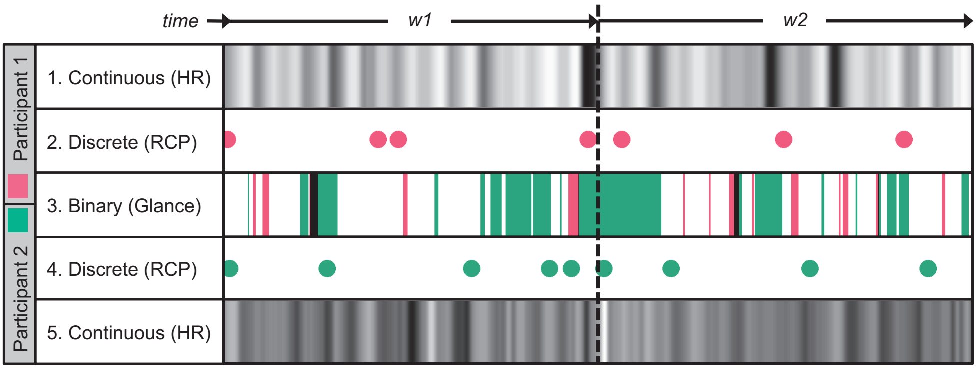


Figure 3.5: Strip plot of individual participant data for a two window (*w1* and *w2*) segment of a single improvisation task (under visual-contact condition). Rows 1 and 5 show continuous heart rate data (lighter colours represent higher values). Rows 2 and 4 represent discrete rhythmic change points (RCPs). Row 3 shows glance, which is binary data (glancing or not-glancing) shaded according to the participant. Black shading represents a mutual glance.

each other. Figure 3.6 shows an example of the extracted extrema and RCPs, which in this case appear to be aligned to heart rate minima.

The null hypothesis was that the number of aligned points would not be significantly greater than the chance outcome. In order to estimate the likelihood of a given number of alignments, each set of RCPs was considered as a randomly distributed set of discrete time points; each of which had a probability of being aligned with an extrema point, defined by:

$$P(aligned) = \frac{n_e(W/res)}{T_s/res} \tag{3.1}$$

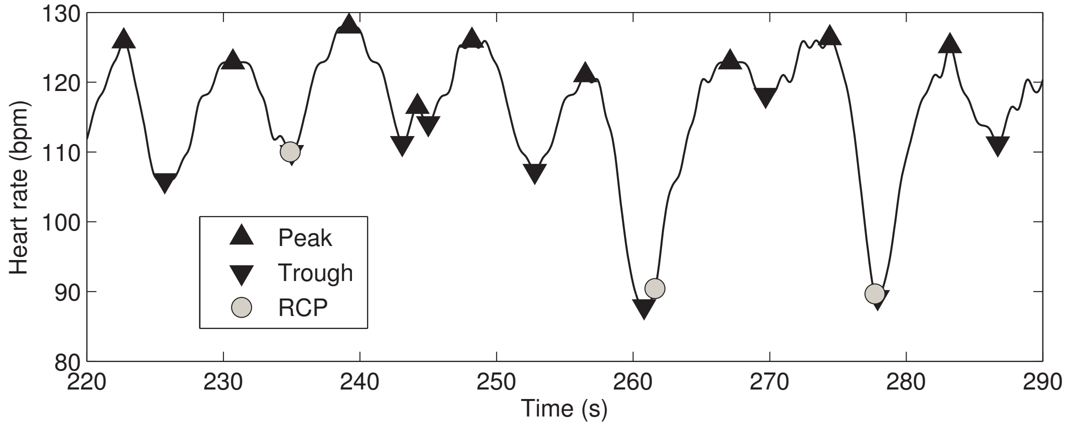


Figure 3.6: Plot of a segment of heart rate data showing detected extrema and corresponding timing of rhythmic change points (RCPs).

The numerator in Equation 3.1 represents the total number of time points on which a RCP could be considered aligned; where n_e is the number of extrema, W is the window size, and res is the time resolution (seconds between samples). The denominator represents the total number of time points, where T_s is the length of the task in seconds. The window size and resolution were 0.5 and 0.1 seconds respectively. These values were chosen based upon the precision of the heart rate data and expected error due to RCP annotation inaccuracies. The p-values were subsequently calculated using the complementary cumulative distribution function (CCDF):

$$p = 1 - F(n_a | n_r, P(aligned)) \tag{3.2}$$

Where F is the binomial CDF, Equation 3.2 gives the one-tailed probability of obtaining *at least* n_a aligned points, given the total number of RCPs (n_r) and the probability ($P(aligned)$) of a single alignment. This is a suitable statistical test for the null hypothesis, since the number of aligned points was considered to be randomly distributed, with a known probability distribution. Table 3.2 presents the results of this analysis. Four of the participants show significant alignment percentages, where:

$$\%_a = 2(n_a)/(n_r + n_e) \tag{3.3}$$

For each of the participants with significant alignments in Table 3.2, average heart rate time series were plotted over the 7 seconds preceding and following all of their RCPs. This analysis was performed to facilitate a visual inspection of how individual participants' heart rates changed on average leading up to, and preceding RCPs. This method was successfully used by Egermann et al. (2013) and Ravaja et al. (2006) to reveal time-windowed patterns in event-based physiological responses. The results are shown in Fig. 3.7. For D1.p1.NV, heart rate rises in the pre-RCP phase and peaks around 2 seconds after the RCP. This moderately corresponds with the results in Table 3.2, which show a significant percentage of maxima alignments for this participant. The strength of the correspondence may be diluted by the fact that the HR time series were averaged across all RCPs, and this participant also had a significant alignment percentage for all extrema. For D1.p2.V the changes in HR are also subtle, however, the mean HR does appear to rise prior to the RCP and drop immediately after. This supports the significant results for maxima alignments in the visual condition. For D3.p2.V the mean HR peaks around 1s prior to the RCP, and troughs around 3s after the RCP. This does not correspond well with the results in Table 3.2. Again, this could be a result of the averaging effect, since the results in Table 3.2 show that this par-

Table 3.2: Percentages of rhythmic change points (RCPs) and heart rate extrema that are aligned for individual participants (Par) within each condition (Cond).

Dyad	Par	Cond	% Aligned RCPs and Extrema ($\%_a$)		
			all extrema	**minima**	**maxima**
1	1	NV	**13.5*****	7.4	**14.8*****
1	2	NV	8.6	5.8	9.3
2	1	NV	4.9	4.4	4.4
3	1	NV	4.5	2.1	6.2
3	2	NV	7.2	5.3	7.9
5	1	NV	5.0	**8.9****	0.0
5	2	NV	5.6	7.1	2.8
1	1	V	6.6	6.8	4.5
1	2	V	7.3	1.4	**11.5***
2	1	V	4.0	2.4	4.9
2	2	V	7.5	6.9	7.0
3	1	V	4.5	2.7	5.4
3	2	V	**11.9****	**13.6****	6.9
5	1	V	2.4	4.4	0.0
5	2	V	1.9	1.2	2.3

Note: Results of one-tailed cumulative CDF: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Data is absent for D4, and D2.p2.NV due to sensor failure.

ticipant also had significant alignments with all extrema. For D5.p1.NV a substantial drop in HR occurs during the pre-RCP phase, followed by an immediate rise during the post-RCP phase. This strongly supports the significant result for non-visual minima alignments presented in Table 3.2. In this case the participant had no corresponding maxima alignments in that condition, which might explain the pronounced minima alignment in the plot.

3.4.2 Dyad-Level Analyses

Correlations Between Self-report Scores

The investigation of relationships between participants' self-report scores within dyads was of particular interest. The justification for this was that it provides an indication of the level of within-dyad agreement with respect to the subjective evaluation of collaborative music making experiences. Windowed SR data were used in order to perform one-tailed, pairwise Spearman correlation analysis between sets of equivalent SR scores within dyads. Spearman correlation is an appropriate test because normally distributed variables and the existence of linear correlations cannot be assumed. The results are shown in Table 3.3, along with the mean correlation across dyads for each SR item. The majority of correlations are weak, indicating that participants did not

Table 3.3: Spearman correlation coefficient (r_S) values for significant within-dyad correlations between self-report ratings

Data	**Feature**	**r_S within dyads (df)**					**Mean**
		D1 (10)	D2 (10)	D3 (8)	D4 (10)	D5 (10)	
SR	Creativity	0.28	0.37	−0.24	0.29	0.02	0.15
-	Engagement	0.41	−0.55	0.42	0.33	−0.12	0.10
-	Energy	−0.07	0.08	−0.11	**0.54***	0.35	0.16
-	Positivity	**0.56***	0.11	0.00	0.12	0.12	0.18
-	Boredom	−0.38	0.18	0.15	0.49	−0.27	0.04
-	Leadership	−0.09	−0.49	**−0.72****	**−0.54***	−0.46	−0.46

Note: Results of Spearman's rho: $^{*}p < .05$, $^{**}p < .01$. df = degrees of freedom.

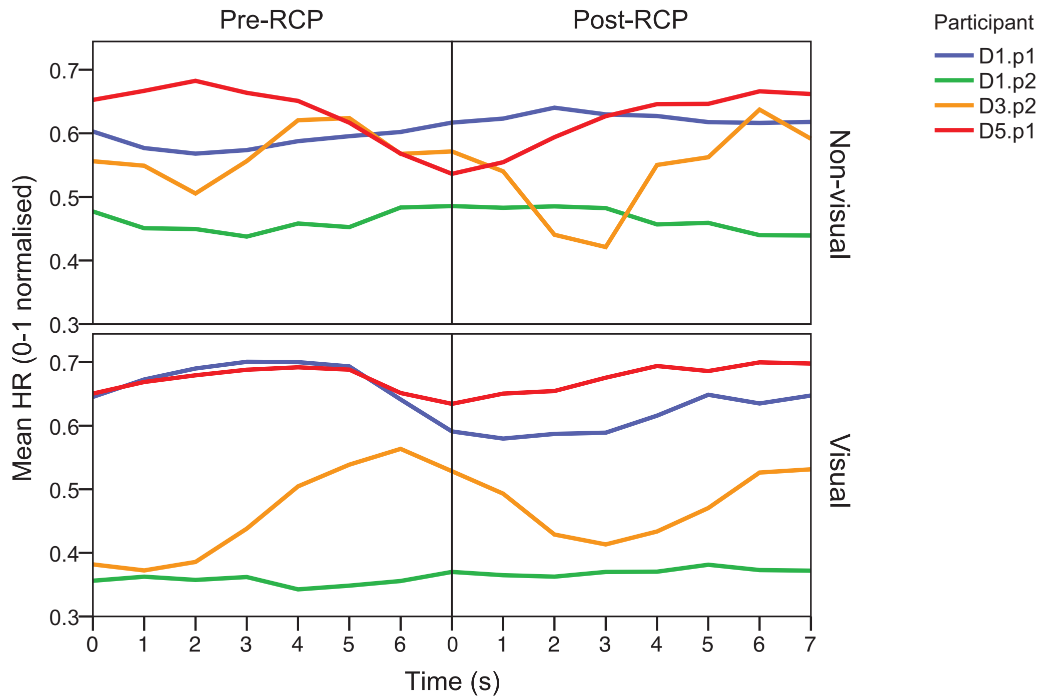


Figure 3.7: Plots of the mean heart rates (HR) during the 7 seconds before and after rhythmic change points (RCPs) for participants with significant RCP-extrema alignments.

generally agree with each other on the subjectively rated aspects of their performances. Correlations for the item *leadership* are strongest (mean $r_S = -0.46$), with significant results for dyad 3 ($r_S(8) = -0.72$, $p < .01$) and dyad 4 ($r_S(10) = -0.54$, $p < .05$).

Heart Rate Synchrony

For the HR measurements, a more focused analysis was undertaken of correlations between the continuous time series data for each participant across entire *Improv* tasks. This was inspired by the studies on emotional contagion and physiological linkage reviewed in Section 2; which suggest that aspects of co-present experiences might be apparent in physiological measurements. A starting point for the analysis was the investigation of whether synchronised changes in HR occurred within dyads; since this would serve as a basic indication of potential physiological linkage. The requirement for synchrony is that a change in one participant's HR should be accompanied by a corresponding change in the other participant's HR. Therefore, to carry out this analysis, each participant's continuous HR data were differentiated and within-dyad comparisons of the slope at each time point were performed. If, at a particular time point, two participants' heart rates shared the same direction of change (positive or negative slope) then that time point was labelled as 'in-sync'. It was then possible to calculate the proportion of synchrony between participants. The results are shown in Table 3.4, where it can be seen that the values all lie close to the chance value of 50%; since 0% and 100% correspond to perfect negative and positive synchrony respectively. This suggests that there was no synchrony between HR changes within dyads. Consequently, the decision was made not to explore a more detailed analysis of heart rate synchrony.

Table 3.4: Synchrony in continuous heart rate data between participants, within dyads. Expressed as the percentage of time where both participants' heart rates were changing in the same direction (positive or negative slope).

Condition	**% Synchrony**				
	D1	**D2**	**D3**	**D5**	**Mean**
Non-visual	46	44	51	53	49
Visual	52	52	49	55	52
Mean	49	48	50	54	50

3.4.3 Study-Level Analyses

In order to perform analyses at the level of the entire study, it was necessary to take into account the nested structure of the study. Each windowed data point comes from an individual, nested within one of the two conditions. These individuals are, in turn, nested within dyads, which are nested under the entire study group. A popular approach to the analysis of data that is structured over multiple levels is the use of linear mixed models (LMMs). Numerous studies have adopted LMMs for the analysis of continuous and subjective human measurements (Glowinski et al., 2013; Egermann et al., 2013; Baayen et al., 2008; Ravaja et al., 2006). Linear mixed models estimate the relationship between a dependent variable and associated covariates by taking into account both fixed and random effects. They also allow for missing data points for subjects (West et al., 2007), which was the case with the data-set in this study. For a single subject i the general LMM specification is given in Equation 3.4 (West et al., 2007).

$$Y_i = \underbrace{X_i\beta}_{\text{Fixed}} + \underbrace{Z_i u_i + \varepsilon_i}_{\text{Random}} \tag{3.4}$$

The vector Y_i represents continuous observations of the dependent variable for subject i. X_i is an $n_i \times p$ design matrix, containing the measured values of the p covariates for each n_i observation collected on the i-th subject. β contains the fixed-effect parameters, which are represented by a vector of p unknown regression coefficients. Much like X_i, Z_i is a matrix representing observed values of the dependent variable. However this matrix is multiplied by u_i, which is a vector of random variables representing the random effects associated with the covariates in Z_i. The final term, ε_i, in Equation 3.4 is a vector containing the residual error terms for each observation of the covariate.

The `MIXED` procedure in SPSS (Version 22) statistical analysis software was used to model the relationships between pairs of continuous features. For the random effects, random slopes and intercepts for participants were specified. Random effects were not included at the level of the dyad, since dyad association did not generally have systematic effects on the individual participant responses. The analysis was carried out with data across both conditions, as well as within conditions. For the former, the random effects model specified that the covariates were nested within the condition. The restricted maximum likelihood (REML) method was used for parameter estimation. To select the best covariance structure for each model (either *variance components*, or *unstructured*) likelihood ratio tests were used to compare the -2 restricted log likelihood values for the models obtained using each of the two covariance structures. If the two models were not significantly different ($p < .05$) then the simpler covariance structure (variance components) was selected. The covariance structure used for each analysis is specified in the notes below the corresponding results table.

Correlations between Continuous Features

Much like the individual-level analysis in Section 3.4.1, analysis was conducted on the way in which glance and RCP events related to changes in other continuous measures. Table 3.5 presents the results obtained from the LMM analysis (see Appendix A.4 for more details). Since there are many potential combinations of the 28 continuous features, only the significant results are presented. All the results come from the analysis within the visual condition. The mean absolute inter-beat lag is significantly negatively correlated with the number of individual glances ($t(29) = -2.37$, $p < .05$) and mutual glances ($t(27) = -2.79$, $p < .01$). Mean HR is significantly negatively correlated with the number of RCPs ($t(26) = -2.19$, $p < .05$). The first two results are closely related, since there is a collinearity between the number of mutual glances and the number of individual glances. This is due to the fact that mutual glances occur when the individual glances of two participants within a dyad are concurrent.

Table 3.5: Linear mixed effects modelling (LMM) estimates of fixed effects between continuous features.

Parameter	**Dependent Variable**	**Estimates of Fixed Effects**			
		Estimate	Std. Error	df	t
Visual:					
No. of glances	Mean abs. inter-beat lag	−0.0006	0.0002	28.8290	**−2.3700***
No. of mutual glances	Mean abs. inter-beat lag	−0.0011	0.0004	27.1060	**−2.7850****
No. of RCPs	Mean HR	−0.9204	0.4202	26.2340	**−2.1900***

Note: t = t-value, df = degrees of freedom. Significance estimates: $^{*}p < .05$, $^{**}p < .01$. All estimates calculated using *variance components* (VC) covariance type.

In a similar manner to the individual level analysis in Fig. 3.7, a plot was created of the study-level mean heart rate during the 7 seconds preceding, and following RCPs. This was calculated as the mean of the set of all the 15 second normalised HR recordings centred upon every corresponding RCP. The results can be seen in Fig. 3.8, which shows that, in the 7 seconds following rhythmic change (post-RCP), there is a fairly linear average increase in HR. For the pre-RCP phase, a slight rise and fall in HR can be seen in the 5 seconds leading up to a RCP.

Correlations between Continuous Features and Self-report Scores

Using the same LMM procedure as above (see Appendix A.4), an analysis of relationships between continuous features and self-report scores was performed. Again, the analysis was performed across both conditions, and within conditions. The aim was to explore which continuous measures might be suitable as predictors of subjective experience. Significant ($p < 0.05$) correlations are presented in Table 3.6. As can be seen, across the entire analysis all of the SR items have at least one associated continuous measure that gives a significant correlation. The mean body quantity of motion (QoM) is positively correlated with creativity ($t(37) = 3.59$, $p < .001$), engagement ($t(31) = 3.01$, $p < .01$), and energy ($t(33) = 3.16$, $p < .01$). There are significant correlations between the beta-band EEG features and leadership. Again, these correlations are strongest in the NV condition, with leadership correlating with mean L-Beta power ($t(21) = 3.33$, $p < .01$), and mean H-Beta power ($t(34) = 3.25$, $p < .01$). In the NV condition mean H-Beta power is also positively correlated with positivity ($t(22) = 2.70$, $p < .05$), and negatively correlated with boredom ($t(21) = -2.30$, $p < .05$). Leadership is also strongly correlated with the number of beats played, with the strongest

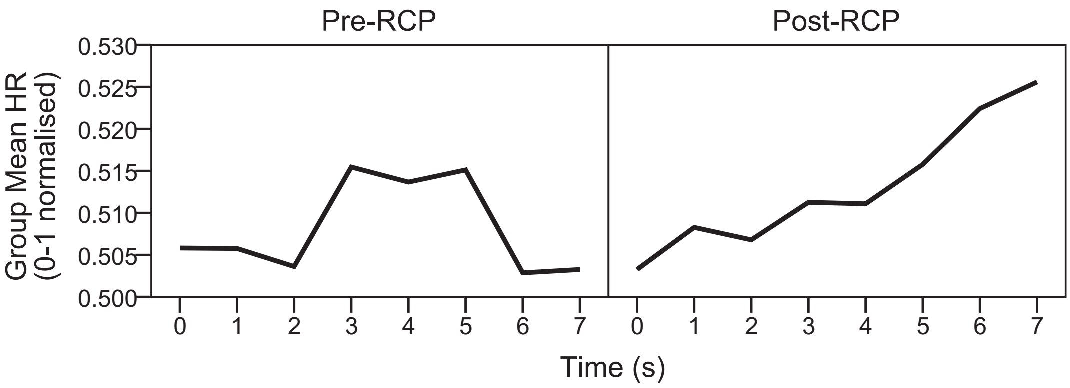


Figure 3.8: Plot of the mean heart rate (HR) for all participants during the 7 seconds before (Pre-RCP), and after (Post-RCP) all the RCPs from both conditions.

Table 3.6: Linear mixed effects modelling (LMM) estimates of fixed effects between self-report measures and continuous features.

Parameter	**Dependent Variable**	**Estimates of Fixed Effects**			
		Estimate	Std. Error	df	t
Both conditions:					
Creativity	Mean body QoM	0.10	0.04	68.28	2.44*
Engagement	Mean body QoM	0.10	0.04	57.67	2.41*
Energy	Mean body QoM	0.12	0.05	54.33	2.56*
Leadership	Mean L-Beta power	0.35	0.17	24.03	2.12*
-	Mean H-Beta power	0.49	0.16	31.53	3.15**
-	No. of beats per sec.	0.17	0.06	83.45	2.75**
Non-visual:					
Creativity	Mean body QoM†	0.21	0.06	37.01	3.59***
Engagement	Mean body QoM	0.17	0.06	30.52	3.01**
Energy	Mean body QoM	0.23	0.07	33.34	3.16**
Positivity	Mean H-Beta power	0.66	0.25	21.64	2.70*
Boredom	Mean H-Beta power	−0.49	0.21	21.34	−2.30*
Leadership	Mean L-Beta power	0.63	0.19	20.86	3.33**
-	Mean H-Beta power	0.81	0.25	34.05	3.25**
-	No. of beats per sec.	0.18	0.08	20.05	2.37*
Visual:					
Energy	Mean body QoM	0.15	0.06	15.85	2.44*
Engagement	Mean velocity	−4.14	1.56	35.99	−2.66*
Boredom	No. of SCR extrema	−0.02	0.01	27.10	−2.77**
-	% Glance time	6.60	3.14	26.16	2.10*
Leadership	No. of beats per sec.	0.21	0.10	41.27	2.06**

Note: t = t-value, df = degrees of freedom. Significance estimates: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Estimates calculated using *variance components* (VC) covariance type or †*unstructured* covariance type.

correlation in the cross-condition analysis ($t(83) = 2.75$, $p < .01$).

The visual condition analyses result in the fewest correlations, with three results that are exclusive to this condition. Firstly, engagement is negatively correlated with the mean MIDI velocity ($t(36) = 2.66$, $p < .05$). Secondly, there is a significant correlation between percentage glance time and boredom ($t(26) = 2.10$, $p < .05$). Finally, the only GSR-related correlation arises here, with a significant negative relationship between the number of SCR extrema and boredom ($t(27) = -2.77$, $p < .01$).

Effects of the Visibility Condition

The final study-level analyses were carried out to evaluate the potential effects of the visibility condition on both subjectively reported and continuously measured aspects of the improvised performances. A similar LMM approach was used to that adopted above, with visibility condition set as a factor, and random slopes and intercepts for participants (see Appendix A.4). Significant ($p < .05$) and noteworthy ($p < .15$) results are shown in Table 3.7, where it can be seen that the only significant effects of visibility condition are on the creativity SR item ($t(9) = -2.26$, $p < .05$). The negative t-value indicates that participants felt more creative in the visual condition, than the non-visual condition.

3.5 Discussion

This study provides three distinct contributions to research on collaborative music making and the use of affective and behavioural sensors in studies of human interactions. Firstly, the techniques, challenges and issues associated with the novel study *design* serve as a valuable reference for researchers planning to conduct similar studies. In addition to this, the *findings* pose interesting research questions, which relate to existing concepts, and serve as inspiration for future studies. Finally, the assessment of various *measures* will assist researchers in deciding which measures and associated sensor technologies are best suited to their particular studies. The following sections discuss these three contributions of the study in more detail.

3.5.1 Design

This study was designed to assess the challenges and issues associated with the experimental use of behavioural and affective sensors as a means of investigating collaborative music making. The conflict between statistical and ecological validity is a noteworthy challenge for researchers attempting to undertake such studies. In this study it was not

Table 3.7: Linear mixed effects modelling (LMM) estimates of fixed effects of visibility condition on all self-report measures and noteworthy ($p < .15$) continuous measures.

Data	**Feature**	**Estimates of Fixed Effects of Visibility Condition**			
		Estimate	Std. Error	df	t
SR	Creativity	−0.66	0.29	9.27	**−2.26***
-	Engagement	−0.89	0.55	17.79	−1.64
-	Boredom	0.44	0.26	9.35	1.69
ECG	No of HR extrema	−0.02	0.01	60.39	−1.63
GSR	No of SCR extrema	−0.02	0.01	55.00	−1.60
MIDI	No of RCPs	0.41	0.26	95.08	1.54
Motion	Mean body QoM	−0.38	0.21	7.07	−1.76
-	Mean arm QoM	0.49	0.25	4.88	1.94

Note: t = t-value, df = degrees of freedom. Significance estimates: $^{*}p < .05$. All estimates calculated using *variance components* (VC) covariance type.

possible to account for the ways in which the experimental environment and measuring devices may have influenced the participants. Chapter 2 highlighted how researchers investigating musical improvisation and ensemble interactions have chosen to tolerate unnatural and highly controlled experimental settings, such as the confines of an MRI scanner; and abstract tasks, such as finger tapping. A one-handed improvised drumming task was chosen for the present study because it was thought to be representative of a basic on-the-fly musical collaboration, whilst also satisfying practical requirements. However, the specificity of this task means that caution must be taken in generalising the findings to other forms of musical interaction.

An additional issue, specific to the design of this study, was the ordering of the visual and non-visual conditions. The decision to use a fixed order was made partly because randomised ordering would not have been effective with a small study size. In addition to this, holding the V condition first would have allowed participants to use non-verbal communication to appraise aspects of their performance. These appraisals could then have had a noteworthy influence on how the participants performed in the second improvisation. As a consequence of this decision, the results may have been influenced by ordering effects. The potential presence of these effects is highlighted when discussing specific findings below.

3.5.2 Findings

The level at which various analyses should be performed is an important consideration for researchers investigating dyadic interactions. It depends both on the nature of the data, and on the questions posed by the researcher. The broad analyses in the present study were performed at the level of the individual, dyad, and entire study. Individual-level analyses investigated how specific events in the performances related to continuous measures. This revealed correlations that may not have emerged during study-level analyses, due to cross-participant random effects such as personality factors and musical ability. Conversely, study-level correlations between subjective and continuous features may not have achieved significance at the level of the individual, due to small sample sizes. The findings are discussed in more detail below, according to their level of analysis.

Individual-level

At the level of the individual, the visual analysis of the continuous data resulted in a more detailed analysis of potential relationships between HR extrema and RCPs. The results in Table 3.2 and Fig. 3.7 indicate that the extent of these relationships varies a great deal between participants. For the majority of participants the relationship is weak or non-existent. However, for some participants it is significant. A plausible explanation for these alignments is psychophysiological linkage. The reasoning behind this is that heart rate is linked to arousal and stress, and participants are most likely to change rhythm when their arousal is too high (e.g. the rhythm is too challenging) or too low (e.g. the rhythm has become boring). Another interesting consideration is that HR and RCP alignments are a representation of the participant attempting to maintain a state of flow, whereby they never become overly challenged, or overly bored by the task. The alignments for D5.p1 are especially interesting, as they appear to correspond solely with HR minima. During an informal post-performance interview this participant commented that she had been particularly nervous about playing in the presence of the other participant, who she knew well. She also had 11 years less drumming experience than her collaborator. It is possible that this anxiety could have contributed to a pronounced rise in HR following new rhythmic contributions. This lends support to the suggestion that HR and RCP alignments were a result of psychophysiological linkage. Furthermore, it highlights the possibility that other factors, such as the participants' personalities, expertise, and relationships, may have influenced these findings. A more focused study is required in order to verify whether a causal relationship exists between HR extrema and musical decisions. In particular, it would be worth considering whether there are other performance-related events that also coincide with HR extrema (e.g. changes in the musical dynamics or specific non-verbal exchanges between musicians).

Dyad-level

The analysis of correlations between SR scores at the level of the dyad indicated that participants did not generally agree upon aspects of their experiences, and the creativity of the performances. These results are similar to those of Schober and Spiro (2014), who found that two jazz musicians did not have a high level of agreement on self-reported evaluations of their performances, despite the performances being judged to be successful improvisations. The most significant and consistent correlations were for the *leadership* item. This is a potentially interesting result, since it suggests that dyads were more attuned to the functional leader-follower aspects of their interaction than they were to the affective aspects, such as energy and positivity. It has been suggested that the establishment and communication of leader-follower relationships is an important tool in the creation of dynamic musical collaborations (Reidsma et al., 2014).

The analysis of HR synchrony within dyads suggested that synchrony did not occur for participants in this study. The existence of physiological linkage between co-present individuals is understudied and previous evidence has been found only in resting participants (Reed et al., 2013). In this study, the active nature of the drumming task may have been the dominant factor in determining changes in HR over time.

Study-level

The study-level analyses of relationships between paired sets of continuous features revealed significant correlations between glance-based features and the mean absolute participant inter-beat lag. These results suggest that glancing at the other participant increases the amount of timing synchrony (decreased lag) between participants. This agrees with existing research, which has shown that visual contact can have a positive influence on timing synchrony between musicians (Vera et al., 2013). A significant negative correlation was also found between the number of RCPs and the mean HR ($t(26) = -2.19$, $p < .05$). This supports the suggestion that RCPs occur at points of low physiological arousal (see Section 3.5.2). If this is the case, then the arousal would be expected to increase following rhythmic change. Evidence for this can be seen in the plot in Fig. 3.8, which shows an average increase in HR in the 7 seconds following rhythmic change.

The study-level analyses of correlations between continuous features and self-report (SR) scores revealed that body motion is positively correlated with *creativity* ($t(37) = 3.59$, $p < .001$), *engagement* ($t(31) = 3.01$, $p < .01$), and *energy* ($t(33) = 3.16$, $p < .01$). This aligns with previous research, which found that quantity of movement features can discriminate between low and high arousal emotions (Castellano et al., 2007). The fact that body movement is also correlated with *creativity* lends support to the *dual pathway model* discussed in Section 2.2.2, which highlights the importance of arousal in the generation of creative ideas. H-Beta EEG power is significantly correlated with *leadership* in both the cross-condition ($t(32) = 3.15$, $p < .01$) and NV analyses ($t(34) = 3.25$, $p < .01$). H-beta correlations were also found with *positivity* ($t(22) = 2.70$, $p < .05$); whilst negative H-beta correlations were found with *boredom* ($t(21) = -2.30$, $p < .05$). Previous studies have associated beta activity with *engagement* and *cognitive challenge* (Budzynski et al., 2008). This concurs with the findings in this study, since one would expect a musician to be more engaged and cognitively aroused when leading the performance. Within the visual condition, the negative correlation between *engagement* and mean velocity ($t(36) = -2.66$, $p < .05$) is interesting, since it suggests that participants hit the drum more softly during periods of high engagement. This might be reflective of participants playing more quietly in order to direct their attention towards the other musician's playing. This result is consistent with descriptions of jazz musicians' multiple levels of attention, incorporating an awareness of both the self and the other (Sawyer, 2003). Also of interest is the correlation between percentage glance time and *boredom*, which suggests that prolonged glances might be used within the performance to communicate boredom; or that they might be interpreted as indicators of boredom when reviewing the video recordings of the performance. The same can be said for the negative correlation between *boredom* and the number of SCR extrema, since SCR events have been shown to be concomitant with emotional arousal (Benedek and Kaernbach, 2010).

The lack of significant correlations between *creativity* and EEG features is interesting because it may be indicative of the use of contrasting thought processes during creative action. This concurs with previous EEG research, which has struggled to show conclusive links between creativity and localised brain activity (Dietrich and Kanso, 2010). Indeed, the literature on creativity often refers to the roles and interplay of different styles of thinking in the development and generation of ideas. For example, the terms *associative* and *analytic* have also been used to define two contrasting modes of thought involved in the creative process (Gabora, 2002; Simonton, 1975).

During the study-level analyses of correlations between continuous features and SR scores, attention was drawn to the fact that the results for the NV condition yielded more, and stronger correlations than those for the V condition. A potential explanation for this is that participants' non-verbal behaviours and physiological responses are more closely tied to their personal subjective experiences when they are unable to see the other musician. This idea is influenced by the theories of emotional contagion and mimicry (see Section 2.2.1), which suggest that being able to see the other participant might influence changes in behaviour and physiology. More work would need to be done to establish whether this is an influencing factor when correlating subjective scores with measures of behaviour and affect.

The results for study-level effects of the visibility condition indicated that the only significant effects of co-visibility were on self-reported creativity ($t(9) = -2.66$, $p < .05$). This indicates that creativity was rated more highly in the V condition. However, it is necessary to be aware of the potential influence of ordering effects on this result. Given the small size of this study, noteworthy results ($p < .15$) were also reported. These indicate that self-reported engagement was generally lower in the NV condition ($t(18) = -1.64$), whilst boredom was higher ($t(9) = 1.69$). With regard to ordering effects, one might have expected boredom to be rated more highly in the second condition

(V). These results support the idea that co-visibility is influential to the subjective experience of collaborating musicians.

The HR and SCR extrema features are the only physiological features that show noteworthy visibility condition effects. These features represent the number of phasic shifts in a participant's physiological arousal over time. The results suggest that there are more shifts during the V condition. This may be reflective of the fact that the participants had more to attend to when they were able to see their fellow musician. This could have led to them shifting their attention more frequently, resulting in corresponding shifts in physiological arousal due to psychophysiological linkage.

The investigations in this study predominantly concerned relationships between pairs of features and conditions. This approach was chosen due to the exploratory nature of the study and the broad array of measures and features collected. A more comprehensive analysis of specific correlations and effects would require the inclusion of other factors, based upon specific hypotheses. Such a multi-factorial approach could yield results that contrast with those presented in the present study. In addition to this, the small sample size used in this study means that the findings should be viewed speculatively in relation to existing studies, and as inspiration for future research.

3.5.3 Measures

The suitability of measures are assessed based on three factors; *practicality*, *reliability*, and *informativeness*. Practicality concerns how difficult it is to collect the measurements, given the constraints of live musical performance. Reliability addresses the issue of whether the raw measurements are likely to contain reliable information, as opposed to noisy, or artifactual data. Finally, informativeness concerns the amount of useful information that could be obtained by extracting features and analysing the measurements. These factors were chosen in light of study aims (see Section 3.1).

Of the sensor-based measures, ECG and accelerometer data were easy to acquire using wireless sensors, and were not particularly prone to noise. On the contrary, the EEG sensor required a lot of preparation and adjustment prior to use, and the data required manual processing (artifact removal), and were prone to noise. Regarding the informativeness of sensor-derived features, the extraction of heart rate extrema time points was found to be useful for the analysis of time-based performance events. The bodily quantity of motion (QoM) feature appeared to be a good indicator of the arousal dimension of subjective experience. Glance related features were acquired by manually annotating video footage; however, it is feasible that modern eye-tracking sensors could perform this task automatically. Glance count and average duration appear to be promising features for the analysis of collaborative interactions where the musicians are visible to one another.

Performance related features, acquired through MIDI data, assisted in the identification of rhythmic changes, and provided velocity and timing synchrony features. Timing synchrony was extracted from beats that were perceptually synchronous, so would not have accounted for more pronounced and intentional changes in tempo. Reliable measures of tempo changes are difficult to extract automatically, and this contributed to the decision not to perform tempo analysis. However, this is an important aspect of improvisation and would be worth considering in future studies. One issue is that MIDI data can only be obtained from certain instruments. Alternative methods, such as audio feature-extraction, should be considered for more generalised studies of collaborative music making.

Finally, the SR measures were obtained using a post-performance questionnaire (PPQ). The PPQ was designed to be as concise as possible, due to the fact that each participant had to complete it from 6 up to 10 times. Instead of including multiple questions addressing similar constructs, specific questions were selected and adapted from existing questionnaires used in related studies. Verifying the validity and reliability of questionnaires is a challenge for any researcher attempting to collect subjective measures of collaborative music making. In particular, reliability is difficult to ascertain, given that the unique nature of each collaboration prohibits the collection of repeated measures. Furthermore, the value and general reliability of introspective SR data has been called into question (Pronin, 2009). These issues provide weight to the argument that alternative measures, such as physiology, might be more suitable tools for the measurement of behaviour and affect. The predicament facing researchers is that, in order to interpret and understand how such measurements relate to affective and behavioural phenomena, existing knowledge is required about the phenomena being measured. This is especially difficult when it comes to investigating collaborative music making: an activity that involves a complex tapestry of contextual and subjective information that changes continuously as the interaction unfolds.

In summary, researchers should take caution in drawing links between self-report and continuous physiological and behavioural data. In particular, it is necessary to carefully consider the causal mechanisms and pathways that might explain such links.

3.6 Summary

The study reported in this chapter set out to i) **assess** the use of behavioural and affective sensors for the investigation of collaborative music making; ii) **report** exploratory findings; and iii) to **identify** suitable measures and features. To address these aims an exploratory approach was adopted, involving the collection and analyses of self-report and *continuous* behavioural and physiological data from pairs of improvising percussionists. The findings indicated that self-reported measures of creativity, engagement, and energy were correlated with body motion; whilst EEG beta-band activity was correlated with self-reported positivity and leadership. Relationships were also observed between cardiac activity and performance events; with heart rate extrema appearing to coincide with rhythmic change points for some participants. Regarding co-visibility, lack of visual contact between musicians had a negative influence on self-reported creativity. The number of glances between musicians was positively correlated with rhythmic synchrony, and the average length of glances was correlated with self-reported boredom.

In view of these findings and the insights gained from collecting and analysing these data, individual measures were assessed according to their reliability, practicality, and informativeness. This led to the identification of **cardiac activity**, **body motion**, and **glance** as particularly promising measures for the investigation of collaborative music making. In the next chapter these measures are considered and reviewed in detail; which subsequently informs the design and development of a prototype device to provide musicians with real-time feedback about the gaze or body motions of their co-performers.

Chapter 4

The LuminUs

Designing a Device for Collaborating Musicians

The previous chapter reported a study, in which physiological, motion, gaze, performance, and self-report data were collected from pairs of improvising percussionists. Following the exploratory analysis of these data, cardiac activity, body motion, and glance were identified as promising measures for sensing affective and behavioural aspects of collaborative music making. This chapter builds upon those findings and poses the question:

> *How can specific sensors and features be used in the design of a device to enhance affective and behavioural interaction during collaborative music making?*

The chapter begins by outlining **design considerations** concerning the selection of input measures and output modalities. In light of these considerations, an in-depth review of **related work** is then carried out in Section 4.2. Sections 4.3 and 4.4 then describe the design and development of the **LuminUs** - a device to provide collaborating musicians with visual feedback about the gaze and body motions of their co-performers.

4.1 Design Considerations

The second aim of this thesis, set out in Section 1.2, is to develop and test a device that utilises sensor technologies in order to enhance collaborative music making interactions. This section introduces and discusses two major considerations for the design of such a device: **input measures**; and **output modalities**.

4.1.1 Input Measures

In light of the findings reported in the previous chapter, three measures were identified for sensing affective and behavioural aspects of collaborative music making: gaze, motion, and cardiac activity. An overview of these measures, and the findings that led to their selection is provided below:

Gaze: The video footage from Study 1 was used to extract the timings of when participants were looking towards each other. The analysis of these data suggested that gaze was used of a means of communication, and that it also had effects upon timing synchrony. The fact that these data were obtained through manual annotation is not seen as a problem, since numerous devices and techniques exist for the automatic detection of gaze.

Motion: Quantity of motion (QoM) was extracted from accelerometers worn on the participants' heads, bodies, and non-drumming arms. Linear mixed model (LMM) analyses showed significant correlations between body movement and self-reported measures of creativity, engagement, energy, and leadership. These results suggest that body motion could convey important information about the success and quality of collaborative interactions.

Cardiac activity: Heart rate was extracted from ECG measurements. The visual and statistical analyses of these data showed interesting and significant relationships between heart rate changes and the points at which participants made musical decisions (introducing a new rhythm). These findings suggest that cardiac activity could be used to provide musicians with feedback relating to the musical intentions of their co-performers.

An additional factor in the decision to pursue these measures is their practicality, and the potential for them to be utilised for affect sensing in real-world scenarios. It is envisaged that all three measures will be accessible through widely used portable consumer-electronic devices within 5-10 years. Accelerometer-based motion sensing is already incorporated into mobile phones and computer gaming devices. Apple have released a smart-watch, which is able to measure heart rate, as well as motion (Limer, 2014). Whilst Google have patented eye-tracking technology (Neven, 2013), which could be incorporated into a set of everyday glasses. Section 4.2 provides an in-depth review of measurement techniques, features, and relevant work relating to each of the three measures discussed above.

4.1.2 Output Modalities

In order for a sensor-based device to support or enhance collaborative music making, it must provide some form of output. In the field of HCI a range of output modalities are used in the design of human-computer interfaces. These include visual displays, haptics, and auditory outputs (Sears and Jacko, 2007). When selecting output modalities it is necessary to consider multiple factors, such as the type of information to be output, the user requirements, and the usage context (Obrenovic and Starcevic, 2004; Reeves et al., 2004). Taking these considerations into account, the following design requirements were specified for the output modality:

Visual: During collaborative musical performance musicians must continuously attend to their own musical contributions, and to those of their co-performer(s). Therefore, auditory feedback is not well suited to this context. Vibro-tactile actuators have been used in the design of systems for providing performing musicians with real-time haptic feedback (Dalgleish and Spencer, 2015; Grosshauser and Hermann, 2009; van der Linden et al., 2011). However, in each case, this feedback merely served to alert the musicians to issues relating to their playing. When it comes to providing more complex and dynamic feedback, haptics are less suitable. Consequently, the device should use a visual feedback modality.

Minimal: A complex interface could distract from social and communicative aspects of co-present collaborative music making. This would not be a desirable outcome for a device designed to enhance collaborative music making. Consequently, the output modality should provide simple, and minimal feedback.

Dynamic: Affective and behavioural signals are highly dynamic; they vary continuously over time, and humans continuously monitor them during interactions (Vinciarelli et al., 2009). Therefore, the output modality should be capable of representing continuous, dynamic information.

Section 4.2.4 provides an in-depth review of considerations and existing work that relate to the minimal, visual feedback of dynamic information.

4.2 Related Work

The previous section outlined considerations for the design of a sensor-based device to enhance collaborative music making. Three potential input measures were chosen: gaze, motion, and cardiac activity. This section presents a comprehensive review of work relating to these measures. In each case the following topics are addressed: measurement methods, feature extraction, and relevant research. Regarding output modalities, the previous section specified that the output should be visual, minimal, and dynamic. Consequently, the final part of this section discusses related work concerning the provision of minimal, visual feedback of dynamic information.

4.2.1 Gaze

Eye gaze is defined by the movements of a person's eyes over time, in relation to features within their external environment. Studies of eye gaze tend to be highly context specific. For example, studies of the ways in which people use gaze to extract information from paintings (Yarbus, 1967), or the coordination of gaze between two people during a conversation (Cummins, 2012). The research in this thesis predominantly concerns the role that gaze plays in non-verbal communication (NVC) between two or more co-present persons. This section initially discusses the measurement of gaze using eye-tracking technologies, and some common gaze-related features. It then provides a review of relevant research on the functions of gaze in NVC, and the use of eye-tracking in studies involving collaborative interactions.

Measurement Methods

The process of measuring the movement or position of a person's eyes is called eye-tracking. Methods for eye-tracking have existed for more than a decade (Jacob and Karn, 2003), with early devices requiring contact with the surface of the eye in order to detect its movement. Most modern eye-tracking devices are non-invasive, using video-based methods to analyse the movements of one or both eyes. An exception is electrooculography (EOG), which uses electrodes placed on either side (above and bellow, or left and right) of the eye in order to measure changes in electrical potentials relating to the eye's movement. Since this method is not particularly accurate for gaze tracking, this review will focus upon video-based methods.

The anatomical and reflective properties of the eye facilitate a number of different methods for tracking using video-based techniques. A popular method is to use image processing to detect the boundary between the dark pupil at the centre of the eye, and the lighter iris which surrounds it (Fuhl et al., 2015; Bozomitu et al., 2015; Gwon et al., 2013). The centre point of the detected pupil can then be used to calculate the position of the eye relative to the head. This method requires that the head is stationary relative to the video camera, which is normally achieved by mounting the camera on a headset

(Javadi et al., 2015). For desktop or screen mounted eye trackers a more advanced method has to be used to account for head movement. The most common method relies on the fact that when infra-red light is shone upon the eye, various reflections are created, known as Purkinje images (Crane and Steele, 1985). The light reflected off the cornea (outer eye) creates a small reflection called a glint (1st Purkinje image), whilst the light reflected from the pupil creates a darker, disc shaped reflection (4th Purkinje image). When viewed through a stationary video camera, both reflections will move by the same amount during head movement, but will move relative to each other during eye movement. The two movements can be distinguished during image processing, and the direction of gaze can be accurately determined. This has been referred to as the dual-Purkinje-image (DPI) method (Crane and Steele, 1985), or pupil centre corneal reflection (PCCR) technique (Tobii Technology, 2010). Commercial technologies for the accurate measurement of eye gaze are now fairly advanced, featuring small high-definition cameras and the ability to wirelessly stream gaze data for real-time analysis. However, these devices are expensive (~£10,000 for a set of eye-tracking glasses). This means that acquiring multiple eye-tracking devices for studies of collaborative interactions may not be a viable option for many researchers. However, in recent years cheaper options have become available. The Pupil project[1] is a mobile eye-tracking hardware and software platform that spawned from a Master's project at the Massachusetts Institute of Technology. It provides a functional, open-source, and low-budget (~£350 at the time of writing) solution for glasses-based gaze tracking (Kassner et al., 2014).

As a final note, it is important to consider the accuracy, precision, and potential errors associated with eye-tracking technologies. Accuracy and precision have been defined as follows (Kassner et al., 2014; Mulvey, 2010):

Accuracy: This is defined as the average angular offset (in degrees of visual angle) between detected fixation locations and the actual fixation target. However, there is little consensus on the exact definition of a fixation.

Precision: This is defined as the Root Mean Square (RMS) of the angular distance (in degrees of visual angle) between successive samples: (x_i, y_i) to (x_{i+1}, y_{i+1}).

Potential error sources in eye-tracking data can be attributed to human, environmental, and technological factors (Kassner et al., 2014; Tobii Technology, 2010). In addition to this there may be timing and latency issues associated with system latencies, and streaming synchronisation (Kassner et al., 2014). These performance factors

[1] http://pupil-labs.com/pupil/

should be taken into consideration according to the required accuracy, test-environment, and choice of technology.

Feature Extraction

In a comprehensive review of eye-tracking studies in usability research, Jacob and Karn (2003) identify a common set of definitions relating to eye-tracking features used throughout the literature:

Fixation: This is defined as a stable gaze direction, with a specified dispersion threshold (typically ~2°), minimum duration (typically 100-200 ms), and maximum velocity threshold (typically 15-100° per second).

Gaze Duration: This is used to define the cumulative amount of time spent gazing at a particular area of interest. The terms 'dwell time', 'glance', and 'fixation cycle' are also used.

Area of Interest: This is an area within the external environment, or visual display, which the researcher defines as being of particular interest.

Scan Path: This defines the spatio-temporal arrangement of a sequence of fixations.

In addition to these common definitions, Jacob and Karn (2003) identify the six most frequently used features in eye-tracking studies:

1. Number of fixations (overall)
2. Percentage gaze time (on each area of interest)
3. Mean fixation duration (overall)
4. Number of fixations (on each area of interest)
5. Mean gaze duration (on each area of interest)
6. Fixation rate (fixations/s) (overall)

Additional features were identified, which are specific to studies involving comparisons of eye-tracking data between two or more people. These studies tend to involve the use of a shared workspace (screen or image) to perform a collaborative task. A simple metric in such studies is to count the number of fixations upon the areas corresponding to each participant's active contributions to the task. For example, in a study where participants played a co-operative game of Tetris, the number of fixations on a player's own blocks, and upon their partner's blocks were counted (Jermann et al., 2010). A similar metric was applied in a separate study where participants completed a pair programming task in which text selections were shared (Nüssli, 2011). It is also possible to look at the amount of overlap between the gaze density peaks of two participants sharing an identical visual display (Cherubini et al., 2008). In this case the gaze data can be windowed in order to examine a particular segment of the collaboration.

Relevant Research

Functions of gaze in non-verbal communication: Studies have shown that gaze has a variety of functions in non-verbal communication (NVC) (Kendon, 1967; Argyle et al., 1973). These functions revolve around the eye's dual role in both the acquisition and expression of detailed information. This can consist of emotional and interpersonal information such as liking, attentiveness, competence, attraction, and dominance (Kendon, 1967; Kleinke, 1986). In this case, gaze is often closely coupled with facial expressions (Adams and Kleck, 2005), bodily gestures, and conversational utterances (Kendon, 1967). For example, research has shown that during conversation speakers tend to look away from the listener at the beginning of an utterance, and towards the listener at the end of an utterance (Kendon, 1967; Oertel et al., 2012). It has been suggested that the speaker directs their attention towards the listener at the end of a turn in order to obtain feedback (Levine and Sutton-Smith, 1973; Kendon, 1967) and/or to signal that it is the listener's turn to speak (Duncan, 1972). Alternatively, gaze information may be used as a means of directing attention towards shared references (Argyle and Graham, 1976) and surveying activity in the external environment. These functions of gaze are particularly valuable during collaborative group work, allowing collaborators to monitor each others' actions and establish mutual grounding (Chanel et al., 2013). Schutz (1976) notes that visual co-presence during collaborative music making allows each performer to share "in vivid present the Other's stream of conciousness in immediacy" (Schutz, 1976, p. 176). This, he elaborates, facilitates the anticipation of the Other's actions, and interpretation of potential cues or commands. Despite its apparent significance, surprisingly few studies have been conducted on the use of gaze during collaborative music making. Davidson and Good (2002) found that members of a string quartet used "conversations with the eyes" to convey important information while playing. Whilst a study with a string duet found that visual contact had a positive influence on timing synchrony between musicians for certain types of music (Vera et al., 2013).

The sense of being looked at, or looking into someone else's eyes for a prolonged period, can provoke strong sensations of discomfort or arousal. This experience has been supported by numerous quantitative studies, which have linked mutual gaze to elevated physiological responses (Gale et al., 1978; Wicker et al., 2003; Williams and Kleinke, 1993; Nichols and Champness, 1971). The magnitude of these responses can vary according to certain personality traits, such as social anxiety (Wieser et al., 2009). This leads to the consideration of an important factor in the study of gaze - *individual variability.* The majority of gaze studies discussed in this section have sought to uncover findings which are generalisable across the entire population. However, having studied gaze in dyadic conversation, Cummins (2012) notes that there are systematic modulations in gaze activity that are invariant at the level of the individual, but which vary a great deal between individuals. Patterson's (1982) sequential functional model of non-verbal exchange draws attention to the influence that both individual and environmental factors have on non-verbal interactions. The model has three stages: *antecedents*, *pre-interaction mediators*, and the *interaction phase.* Antecedents factors include personal, experiential, relational, and situational influences. Pre-interaction mediators include behavioural predispositions, potential changes in arousal, and cognitive-affective assessment. The interaction phase includes avoidance and escape considerations, compliance, physical attractiveness, and positive/negative evaluations. Kleinke (1986) uses the sequential model as a framework for providing a comprehensive review of existing gaze studies. It serves as a reminder of the complex array of factors which might influence gaze behaviour. These factors must be considered carefully when designing studies of gaze in NVC.

Eye-tracking and collaborative interactions: Eye-tracking technology has been adopted in studies across a range of fields such as cognitive neuroscience, social psychology, linguistics, and user experience research. Numerous studies have investigated how eye-tracking might support computer-supported collaborative work. This often involves providing each collaborator with an on-screen representation of where their partner is currently gazing within a shared workspace or virtual environment (Jermann et al., 2010; Vertegaal, 1999; Bednarik et al., 2011; John et al., 2014; Müller et al., 2012). These studies frequently report that such gaze-feedback can have negative impacts on the interaction, leading to information clutter (John et al., 2014), uncertainty and ambiguities (Müller et al., 2012), and distraction from the task (Jermann et al., 2010). Other studies have sought to use eye-tracking to quantify the success of remote or computer-based collaborative interactions by looking at features such as eye movement coupling (Chanel et al., 2013), gaze overlap (Cherubini et al., 2008), and gaze cross-recurrence (Nüssli and Jermann, 2012; Belenky et al., 2014; Richardson and

Dale, 2005; Uzunosmanolu and Çakir, 2014). These features all serve a similar function, which is to ascertain the amount of similarity between the spatio-temporal aspects of each participant's gaze data. In each case, high similarity was indicative of more successful interactions.

Studies of gaze in social interactions often involve the use of live video feeds or virtual environments in order to approximate real-time face-to-face interactions (Pfeiffer et al., 2013). Few studies have used eye-tracking to facilitate research on co-present human-human interactions and non-verbal communication. In studies of conversation and turn-taking, eye-tracking experiments have predominantly reinforced existing theories (Jokinen et al., 2010; Vertegaal et al., 2001). Other studies have focused on clinical and diagnostic applications, such as the analysis of mutual gaze between children and adults (Ye et al., 2012), and the gaze interactions of individuals with autism spectrum disorders (Sterling et al., 2008) and social anxiety (Wieser et al., 2009). No studies were found that have used eye-tracking to investigate or support the interpersonal and affective functions of gaze during collaborative musical interactions. Music related studies have generally focused on the potential for eye-tracking to be used as a control medium (Hornof, 2014).

4.2.2 Motion

Body motions can occur as purposeful, sub-conscious, or impulsive reflex actions. As an intrinsic feature of non-verbal expression, body motion is closely tied to gesture (Goldin-Meadow, 2000). Body motion also accompanies changes in posture and head movements, which are associated with emotional expression (Kendon, 1967). The findings reported in Chapter 3 indicated that low-level motion features, such as the quantity of body motion, could be sufficient to provide information about the affective and behavioural states of musicians. Consequently, this thesis does not consider spatial representations of body motion, since this would add a level of complexity that is beyond the scope of this research. With this in mind, this section addresses the measurement of body movement and associated kinematic motion features. A review of relevant research is then provided.

Measurement Methods

In Chapter 2 the following methods for measuring body motion were briefly highlighted: marker-based motion tracking, single camera setups, depth sensitive cameras, and worn accelerometers (see Section 2.1.3). This section provides a more in-depth review of measurement methods, with a particular focus on methods that are suited to real-world applications.

Marker-based: Optical marker-based systems usually comprise multiple infra-red cameras, which are used to calculate the three-dimensional positioning of reflective markers within a designated space. Of all the methods discussed in this section, these systems offer the highest accuracy (Bianchi-Berthouze and Kleinsmith, 2015). However, the requirement for multiple cameras, and complex calibration procedures means they are not well suited for deployment in the real-life-settings (Dael et al., 2016).

Electromechanical and electromagnetic: These methods involve the use of active sensors; such as potentiometers, magnetometers, and accelerometers. The sensors are positioned at designated sites on the body, enabling specific limb and joint movements to be recorded (Bianchi-Berthouze and Kleinsmith, 2015). For full body motion capture the sensors tend to be integrated into a suit. The advantages of these methods are that they are portable, accurate, and are not affected by occlusions; whilst the disadvantages are that they can be costly, and are sensitive to magnetic fields (Dael et al., 2016).

Vision-based: Modern image processing techniques enable the extraction of three-dimensional (3D) body motion features using a single low-end camera (Quesada and León, 2011). Depth-sensitive cameras, such as the Microsoft Kinect, are able to recognise and report the 3D coordinates of a person's skeleton joint positions. These options offer the benefits of being low cost, unobtrusive, and mobile (Bianchi-Berthouze and Kleinsmith, 2015). Furthermore, the performance of the Kinect has been shown to compare closely with that of a 3D marker-based system (Clark et al., 2012; Galna et al., 2014). A drawback of vision-based systems is that they are usually sensitive to lighting conditions and occlusions; which means they have difficulty recognising body movement in complex scenes (Dael et al., 2016).

Accelerometer-based: Camera-less motion measurements can be obtained using single accelerometers worn at positions of interest. These normally track acceleration along three axes, with the main limitation being that they cannot provide accurate velocity information. Despite this limitation, relevant studies have adopted accelerometers for the detection of conversing groups (Hung et al., 2014); the measurement of stage fright in musicians (Kusserow et al., 2010); and the tracking of violinist bowing motions (van der Linden et al., 2011). The advantages of accelerometers are that they are easy to set up, and provide accurate and uninterrupted acceleration data. Dedicated accelerometer devices tend to be small, wireless, and capable of streaming data directly to a computer for real-time analysis. A further advantage is that many consumer electronic devices – such as smart phones, smart watches, and activity trackers – already contain accelerometers, from which data can be accessed using custom-built applications.

Feature Extraction

Regarding the processing and analysis of motion data, kinematic features used in existing studies can be grouped into two distinct categories:

Low-level time-derivative features: Velocity, acceleration, and jerk (derivative of acceleration with respect to time) are commonly used kinematic motion features Pollick et al. (2001); Camurri et al. (2004); Kleinsmith and Bianchi-Berthouze (2013). These features can be amalgamated over time windows in order to provide associated statistics such as the mean, variance, maximum, and minimum.

High-level time series features: Referred to as meta-features (Castellano et al., 2007), these involve consideration of the motion data as a continuous, dynamic time series. For example, this might involve looking for specific motion patterns, or for relationships between the time series data from multiple sources. Frequency components of the motion can also be analysed, using Fourier time series analysis (Holstein, 2002).

In the following section more details are provided on the ways in which certain features have been extracted and analysed in studies of body motion.

Relevant Research

Motion and affect: In order to study body motion in isolation from associated emotional cues, such as facial expressions, researchers have adopted point-light displays. These are created by videoing subjects with reflective markers attached to their joints and then manipulating the video footage such that only the markers are visible against a black background. Studies with point light displays have revealed that humans are able to successfully identify emotions from recordings of mimed actions (Walk and Homan, 1984; Pollick et al., 2001) and interpersonal dialogues (Clarke et al., 2005). In the latter case, emotions were still correctly identified when the videos were flipped upside-down. These studies support the notion that basic motion features might be sufficient to communicate affect; independent of more complex gestural shapes and patterns.

Castellano et al. (2007) attempted to separate motion from gesture using an alternative method, in which they asked participants to express different emotions by simply raising and lowering their arms (in the coronal plane). Using video footage of the arm motions they extracted velocity, acceleration, quantity of motion (QoM), fluidity, and contraction index features. They describe QoM as a measure of the total amount of motion; fluidity as a measure of the uniformity of motion; and contraction index as a measure of the degree of contraction and expansion of the body. From the temporal profiles of each feature they then extracted *high-level* features. These included counts of the number of maxima and minima, and measures of the initial and final slopes in the time series data. They used both the low and high-level features to train various classifiers to recognise the performed emotions. QoM-based features were found to perform best, discriminating between high arousal (anger/joy), and low arousal (sadness/pleasure) emotions. However, the recognition rates for specific emotions were low and the authors highlighted the need for the classifiers to account for individual variability.

Numerous other studies have also found evidence for a link between measures of body motion activity and the arousal/activation aspects of emotion (Wallbott, 1998; Dahl and Friberg, 2007; Sanghvi et al., 2011). There has been less success in relating the valence (positive/negative) aspects of emotions to body movement features (Metallinou et al., 2011; Kleinsmith and Bianchi-Berthouze, 2013; Karg et al., 2010; Cowie et al., 2010).

Real-world applications: In recent years researchers have begun to look towards the potential applications of body motion analysis in real-world applications. This has been catalysed by the increased use of motion sensors in computer gaming and mobile phone related applications, such as daily-activity and sleep tracking. The research in this thesis is particularly concerned with applications in the contexts of HCI and the performing arts. Bianchi-Berthouze (2012) found that interacting with computer games through role-related movements, as opposed to hand-held controllers, led to an increased sense of engagement with the game, and a more affective experience. In the discussion of her results, Bianchi-Berthouze suggests that social interaction in collaborative games could be further facilitated by providing each player with a representation of their collaborator's body movements.

In another HCI-related application, Sanghvi et al. (2011) used high-level QoM and contraction-index features to train classifiers for the automatic recognition of engagement in children playing chess with a robot. Their results showed that QoM was the most successful feature for the recognition of engagement with the robot. Won et al. (2014) developed an automated affective computing system that attempted to predict the success of dyadic teacher-learner interactions using body motion measurements from two Kinect cameras. They found that the summed standard deviation of the teacher's head and torso movements was one of the most predictive features.

Volpe (2003) investigated applications for expressive gesture analysis in the context of the performing arts. He introduced Multilayer Integrated Expressive Environments (MIEEs) as a means of developing novel artistic performances; whereby the performing actions can take place in multiple connected spaces (physical or virtual). Each of these spaces can contain subjects and objects that are real (e.g. a human, or lighting fixture); virtual (e.g. an avatar); or mixed (e.g. an avatar with physical robotic arms). Virtual or mixed subjects are of particular relevance, since they are designed to observe and process expressive gestures in order to convey their own expressive content; much like the device envisaged in this chapter. Volpe describes the mapping of input expressive gestures to output expressive content on three levels: i) expressive direct mapping; ii) expressive high level mapping; and iii) expressive mapping monitoring. Expressive direct mapping is a non-dynamic mapping of the input to the output, which involves pre-defined functions or rules. Expressive high level mapping involves reasoning and decision making-processes; whereby the mapping is dynamic (it can change over time). Finally, expressive mapping monitoring serves to evaluate the effectiveness of the aforementioned mapping layers, and can also act upon these layers to modify their functioning accordingly.

Motion and musical performance: Movements during musical performance can be categorised according to four functions: sound-producing; communicative; sound-facilitating; and sound-accompanying movements (Goebl et al., 2014; Dahl et al., 2010; Jensenius et al., 2010). Sound-producing movements are those which are involved in creating or modifying sound (e.g. strumming a guitar). Communicative movements serve to communicate information between performers, or between performers and an audience. Sound-facilitating movements serve to support sound-producing movements (e.g. preparing to hit a piano key by moving one's arm). Sound-accompanying movements are not at all linked to sound production, but are intended to follow features of the sound (e.g. dancing). Some types of motion may fall into more than one category. For example, foot tapping can have a sound-facilitating function (helping the musician keep time), and a communicative function (Jensenius et al., 2010). Motion capture has been used to investigate non-verbal communication between collaborating musicians (Glowinski et al., 2013; Healey et al., 2005) (see Section 2.1.3 on page 32); and to provide musicians with real-time feedback relating to posture (Dalgleish and Spencer, 2015) and violin bowing (van der Linden et al., 2011).

Body motion analysis has also been applied to the study of affect in musicians. Castellano et al. (2008) carried out a single camera-based analysis of body movements in emotionally expressive piano performances. They used QoM of the upper body and velocity of the head as features; from which high-level time series features were then extracted. Unlike the previously reviewed studies, QoM was not found to be strongly related to emotional expression. This was partly attributed to the fact that the pianist's movements were constrained by the characteristics of the instrument, leading to smaller variations in QoM. However, their findings did reveal variations in some head velocity and QoM features between different expressive conditions; suggesting that these features do convey information about emotional expression.

Dahl and Friberg (2007) created silent video clips of three musicians (marimba, bassoon, and saxophone) performing with different emotional intentions, and asked naive human judges to rate the video clips according to four movement cues: amount, speed, fluency, and regularity. Sadness was associated with slow and smooth movements; anger with jerky movements; and happiness with large and somewhat fast movements. Fear was not well conveyed by movement. The authors highlight individual differences in the use of movement cues for expression through body movements. For example, happiness was characterised by fast movements for the bassoon player, and large movements for the saxophonist.

4.2.3 Cardiac Activity

Cardiac activity is defined as physiological activity that directly relates to the functioning of the heart. During Study 1 heart rate (HR) data were collected from drummers using wireless ECG sensors. The analysis of these data suggested some potentially interesting relationships between HR changes and creative decision making (see Sections 3.4.1 and 3.5.2). This section initially discusses a range of methods for cardiac measurement and feature extraction. Relevant studies of cardiac activity are then reviewed; specifically, those which relate to creativity, decision making, and musical performance.

Measurement Methods

The most accurate way of measuring cardiac activity is through an electrocardiogram (ECG), which typically requires the connection of between three and twelve skin electrodes across points on the body. The resulting measurement is a vector representation of the electrical signals that coordinate and trigger the beating of the heart. Detailed information about the behaviour of the heart can be extracted from analysis of the ECG signals; hence its widespread use in medical diagnostics. Another common method for measuring cardiac activity is photoplethysmography (PPG), which involves measuring fluctuations in blood volume in the skin. This is usually achieved by shining light through the skin (at the fingers or ears) onto a sensor that detects variations in the intensity of particular wavelengths that are absorbed by the blood. A disadvantage of most ECG and PPG techniques is that they require the use of a measurement device that must be attached to the subject's skin. However, the recent integration of cardiac sensors into consumer smartwatches (Phan et al., 2015) demonstrates that this can be achieved in an unobtrusive manner.

Sensors have also been developed to measure ECG signals without direct skin contact (Plessey Semiconductors, 2012; Chamadiya et al., 2012; Oehler et al., 2008). These devices rely upon capacitive coupling techniques, which means that they still have to be positioned close to the subject's body; normally in contact with their clothing. Despite this potential limitation, various real-world applications have been demonstrated, such as the integration of sensors into a car seat in order to detect tiredness (Walter et al., 2011).

Truly contactless solutions, which extract heart rate information from video images of a subject's face, have proven to be accurate (Poh et al., 2011; Pursche et al., 2012). Similar to the PPG method discussed above, these work by analysing small colour changes in the subject's skin. The Philips Vital Signs Camera is a commercially available application that allows users to measure their breathing and heart rate using the webcam on their tablet or smartphone. The major disadvantage of these methods is that they currently require the subject to stay completely still in order for the measurement to be successful (Philips Electronics, 2014).

Feature Extraction

The first step in the extraction of cardiac activity features is to detect the timing of individual heart beats. For ECG data the beat timings are derived from the detection of R-waves – prominent peaks in the heart's electrical activity that correspond with individual heart beats. For PPG data, beat timings are derived from changes in blood volume, and are referred to as pulse-peaks.

Heart rate (HR), normally expressed in beats per minute (bpm), is the most commonly extracted feature of cardiac activity. It is obtained by measuring the inter-beat intervals (IBIs) between consecutive heart beats and calculating $60/IBI$. These intervals are called RR-intervals for ECG data, and pulse-peak intervals for PPG data. HR values can then be grouped over specific time windows in order to extract low-level statistical features, and high-level time series features; much like the motion features discussed in Section 4.2.2.

More complex features of cardiac activity can be obtained from analysis of heart rate variability (HRV), which is a measure of the variation in IBIs between successive heart beats. HRV analysis can be separated into time-domain approaches, and frequency domain approaches; both of which provide potentially informative features:

Time-domain features: Bilchick and Berger (2006) describe a number of commonly used time-domain HRV features:

SDNN: This is simply the standard deviation of all IBIs.

SDNN index: For each 5 minute segment of data the SDNN is found, then the mean of these SDNNs is calculated.

SDANN index: For each 5 minute segment of data the mean IBI is found, then the standard deviation of these means is calculated.

RMSSD: This is the root-mean-square of successive differences. Calculated as the square root of the mean of the squared differences between successive IBIs.

pNN50: The percentage of differences between successive IBIs which are greater than 50 ms (can also be calculated with lower thresholds (Mietus et al., 2002), such as 20 ms).

These features are primarily used for clinical diagnostic purposes and have limited applications in psychophysiological research (Berntson et al., 1997). However, time-domain features have been employed in short-term (~5 min) studies of emotion (Kim and André, 2008), flow experiences (Keller et al., 2011), and cognitive challenge (Wood et al., 2002).

Frequency-domain features: Spectral analysis techniques can be used to extract the frequency components of a time-varying signal. In HRV research this normally involves the use of the fast Fourier transform (FFT) and subsequent calculation of the power spectral density (PSD) function. The PSD defines the power of a signal as a function of its frequency components. As the input data for spectral analysis, it is common to take a series of IBIs and re-sample them to provide instantaneous heart rate values sampled at 4 Hz (Giardino et al., 2002; Bilchick and Berger, 2006). Re-sampling is performed because it is preferable to have evenly spaced time samples when performing FFT calculations.

Low frequency (LF) and high frequency (HF) bands are normally extracted from the HRV power spectrum because there is evidence that the power within each band reflects differing contributions from the sub-systems of the autonomic nervous system (ANS) (Bilchick and Berger, 2006; Appelhans and Luecken, 2006; Berntson et al., 1997). These sub-systems comprise the parasympathetic nervous system (PNS), which functions during 'rest and digest' activities; and the sympathetic nervous system (SNS), which prepares the body for the 'fight-or-flight' response. The most common frequency-domain features are:

High frequency (HF) band: This is defined as the band between 0.15 and 0.4 Hz. HF power has been shown to predominantly reflect PNS modulation. It is related to the variations in heart rate that occur as a result of respiratory movements, which are also under the influence of the PNS.

Low frequency (LF) band: This is defined as the band between 0.04 and 0.15 Hz. LF power has been shown to relate to both PNS and SNS modulation, although there is some debate over the relative contributions of each system (Malik et al., 1996).

Ratio of LF to HF (LF/HF): This is believed to reflect the ratio of PNS and SNS activity, often referred to as the 'sympathovagal balance'.

The accuracy of features that can be extracted from cardiac activity measurements is dependent on the chosen measurement technique; and for HRV analysis, it is recommended that ECG data should be collected at a sampling rate of 500-1000 Hz (Berntson et al., 1997).

Relevant Research

Section 2.1.4 provided some examples of studies where cardiac measurements have been used to investigate the psychophysiological aspects of various affective experiences. The findings from Study 1 showed potential relationships between cardiac activity and creative decision making. Consequently, the literature review in this section specifically looks at work that informs an understanding of the ways in which cardiac activity might relate to short-term aspects of creativity and decision making during musical performance.

Creativity: Literature on creativity often refers to the roles and interplay of different 'styles of thinking' in the development and generation of ideas. For example, there are frequent references to the terms *divergent thinking* (exploring multiple ideas to generate creative solutions to a problem), and *convergent thinking* (using existing knowledge to establish the single-best answer to a problem) (Guilford, 1967). Other researchers have suggested similar sets of distinct thought processes in relation to creativity. Suler (1980) discusses the interaction between *primary* and *secondary process* thinking during creativity. He defines primary process thinking as "loose, illogical, and highly subjective ideation", whilst secondary process thinking involves moulding a novel idea or insight "into a context that is socially appropriate to others" Suler (1980, p. 144). The terms *associative* and *analytic* have also been used to define modes of thought involved in the creative process (Gabora, 2002; Mednick, 1962; Simonton, 1975). Associative thought involves the formation of connections between different ideas and concepts, which may not be causally related. Conversely, analytic thought is more focused and logical, leading to presentable solutions.

Despite the differences in terminology, there are many similarities between the various definitions above, suggesting a general consensus that pairs of contrasting thought processes are involved in creative cognition. Similar dual processes may be at work in the creative mind of the improvising musician. Furthermore, it seems reasonable to hypothesise that these distinct thought processes might be accompanied by comparably distinct phases of physiological activity. A similar hypothesis was put forward by Bowers and Keeling (1971), who carried out a study on HRV and creative functioning. They suggested that creative thought involves cognitive shifts between concern for the external environment and more internally oriented thoughts, such as dreaming and imagination. Additionally, they proposed that shifting between internal and external processes might be accompanied by phases of cardiac acceleration and deceleration. Their hypothesis was supported by their experimental results, which showed a relationship between HRV and scores on a creativity test. However, a more recent experiment by Deininger and Loudon (2011) found no relationship between HRV and divergent thinking. Clearly there is a need for more research in this area, as no further studies of cardiac activity and creative thinking could be found.

Decision making: During collaborative improvisation, the process of contributing new material to the ongoing interaction will often involve concious decisions on behalf of the musician. When making these decisions the musician may have various considerations, such as whether the contribution will sound good; how well it will be received by the audience or fellow musicians; and to what extent it will challenge their musical ability. These thoughts may be quite separate from those involved in the creative process. Consequently, it is worth reviewing the ways in which cardiac activity has been related to decision making alone. One highly influential theory that relates physiological processes to decision making is the somatic marker hypothesis (SMH) (Damasio, 1996). The SMH proposes that decision making is informed by emotions and, more specifically, the feelings that arise as a result of emotion-induced changes in one's body-state. This specific type of feeling constitutes a somatic marker (SM). Over time, SMs become associated with our experiences of the world and the outcomes of our past decisions. When we encounter a situation or decision that is accompanied by a negatively associated SM, we will tend to take evasive action, and vice versa (Bechara et al., 2005). Experimental evidence for the SMH has predominantly come from studies involving the Iowa gambling task (IGT). The IGT is designed to simulate the uncertainty and positive/negative outcomes involved in real-life decision making. Participants try to win as much money as possible by making multiple selections from four decks of cards (A, B, C, and D). Most of the cards result in a monetary gain, which is higher for decks A and B, than for C and D. However, an apparently random proportion result in a more substantial monetary loss. Again, this loss is comparatively higher in decks A and B. Each deck actually has fixed probabilities of reward and loss, such that selecting mainly from decks A and B results in a net loss, whilst C and D result in a net gain. Participants are initially unaware of this and the intention is that they should gradually learn to select from the advantageous decks. Researchers can then investigate the influence of SMs on the evolving decision making process.

The majority of SMH studies involving physiological measurement have used skin conductance response as a measure of physiological activation (Dunn et al., 2006). However, a few studies have adopted cardiac measurements. Crone et al. (2004) investigated changes in heart rate while participants played a modified version of the IGT. In this version, choices B and D resulted in a less frequent, but more substantial loss. This served to introduce variation in the size and frequency of detrimental outcomes. The authors highlight that previous studies have shown that heart rate slows preceding a voluntary response (Lacey and Lacey, 1974; Van der Molen et al., 1985); and that slowing is greater when individuals are preparing for an aversive event (Somsen et al., 1983). Their experimental results showed that heart rate slowed preceding decisions with a high risk of loss, and that this slowing only occurred for good performers on the IGT. Additionally, they found that heart rate slowed immediately preceding punishment, and returned to baseline levels following reward. Lee et al. (2010) used the IGT to explore whether heart rate and HRV measures could be used to predict decisions. They found that mean and standard deviation-based features of the inter-beat intervals (IBIs) were significantly different preceding advantageous choices, compared to disadvantageous choices. In particular, participants' heart rates tended to slow down prior to disadvantageous decisions. The authors used the IBI features to train a pattern classification algorithm to predict the success of decisions, resulting in accuracies that were significantly higher than chance. It should be noted that there is some contention over the validity of the SMH, and of the suitability of the IGT as a means of testing it (see Dunn et al. (2006) for a review).

Studer and Clark (2011) looked at heart rate changes during a Roulette gambling task, where the chances of success were explicitly presented, and where half the participants also had an active choice in how much money they gambled on each decision. They noted that cardiac responses to motivational stimuli are normally *bi-phasic*, comprising a deceleration component, and an acceleration component. Studer and Clark extracted and analysed these components separately, for both the period preceding, and following each decision. They found that HR decelerations following decision making were generally less for participants who had an active choice, compared to those with no choice. HR decelerations preceding decisions were found to be greater for decisions involving higher uncertainty. Additionally, pre-decision HR accelerations were increased by the ability to make a choice, and by higher bet amounts.

In an attempt to investigate decision making in more complex and real-life circumstances, Leone et al. (2012) measured the HRs of participants whilst they played games of speed chess. Analysis was performed on 10 second windows, centred upon the execution of each move. The results indicated that a player's own moves were generally preceded by a rise in their HR, which peaked around 2-3 s after their move. Interestingly, correct moves (compared to ones labelled as blunders) were preceded by a short dip (duration 3-4 s) in the player's HR; with a minimum at around 1.5 seconds before the execution of the move. For opponent blunders, there was a higher increase in the other participant's HR in the 2 seconds following the move. The results of this study are of particular relevance to the research in this thesis, since the act of playing a chess game also involves some degree of creative thought, as well as co-present interaction with another person.

Musical performance: The findings in Study 1 led to a specific interest in the relatively short episodes of cardiac activity that accompany musical decisions. However, when musicians are participating in collaborative music making, there are likely to be additional underlying features of the interaction situation that cause changes in cardiac activity. It is useful to have an awareness of other cardiac influences, so that attempts can be made to account for them when analysing cardiac data. Section 2.1.4 briefly discussed the concept of *flow*: the state of being optimally engaged in an activity. It also highlighted a study with pianists, which found relationships between flow states and various physiological measures (de Manzano et al., 2010). In particular, this study found that two cardiac features – HR and LF/HF ratio – were positively related to self-reported flow. No additional psychophysiological studies of flow during musical performance have been reported. However, studies involving non-musical activities have reported similar findings, with flow states being accompanied by reduced HRV (Keller et al., 2011), and increased HR, and LF/HF ratio (Gaggioli et al., 2013). These studies tend to draw attention to the fact that low heart rate and high HRV measures have previously been related to mental workload and cognitive focus (Hjortskov et al., 2004; Berntson et al., 1997).

Researchers have also shown how cardiac features vary in response to other aspects of musical performance. Harmat et al. (2010) recorded ECG measurements from three professional pianists during the recital of a self-selected piece, and an unknown, technically difficult piece. The average HR during the difficult piece was lower than during the self-selected piece. Additionally, the HF component of HRV increased during the first two minutes of the difficult piece, before decreasing in the proceeding 2 minutes. In a larger study, Nakahara et al. (2009) asked thirteen pianists to play a set piece of music under two conditions: expressive, and non-expressive performance. In the former they were instructed to play the piece expressively, whilst in latter they were told to play the piece without emotions. HR and the LF/HF ratio were both found to be significantly higher during the expressive condition.

Regarding the effects of the music itself, research has shown that both heart rate and the LF/HF ratio increase in response to faster tempos (Bernardi et al., 2006), and that emotional aspects of music can induce physiological changes (Kim and André, 2008). These effects would clearly be difficult to predict during an unfolding musical collaboration. However, it might still be possible to account for changes in dynamic properties of the music such as volume and tempo.

4.2.4 Visual Feedback

The previous sections have dealt with the measurement and interpretation of affective and behavioural signals. In this section attention is shifted towards ways in which information acquired from sensors might be fed back to musicians during collaborative interactions. At the beginning of this chapter three basic requirements for this feedback were defined: *visual*, *minimal*, and *dynamic* (see Section 4.1.2). This section begins with a review of existing visual methods of displaying dynamic (time-varying) information in various applications. Subsequently, it focuses upon work relating to the visualisation of affect, such as associations between colours and emotions.

Visualising Dynamic Information

One of the most widely recognised instruments for displaying dynamic information is the gauge. Analogue needle gauges are a familiar sight for most drivers, as they are still commonly used in cars to display information, such as speed, and engine revolutions. They are also found on weighing scales, barometers, and many scientific and medical devices. In the musical domain, sound levels are measured using a needle gauge called the volume-unit (VU) meter; a device originally developed to measure radio and telephone broadcast levels (Chinn et al., 1940). In modern audio devices, sound levels tend to be represented by a peak meter, consisting of a row of lights which light up sequentially in relation to the peak signal level. There is normally a variation in the colour of the lights, with green representing lower audio levels, and red representing the highest levels. Figure 4.1 provides a review of the lighting layouts and colour arrangements for some contemporary audio devices.

When designing a device for visualising dynamic information, it is important to consider the number of time-varying attributes which need to be displayed. This can

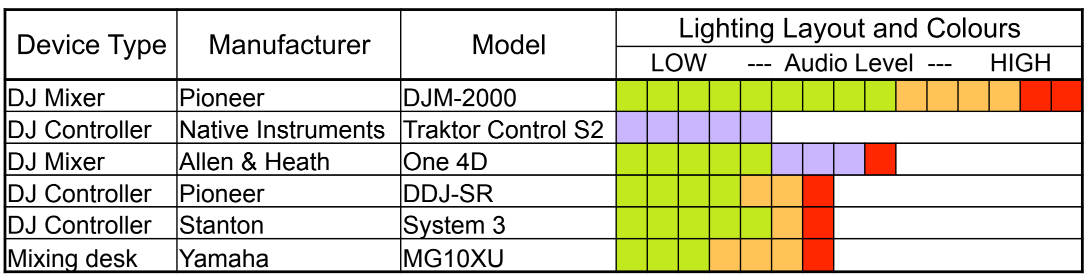


Figure 4.1: A review of audio level displays on contemporary audio devices. Lighting layouts are shown horizontally, however they are vertically orientated on the devices.

then be accounted for in the number of representational degrees of freedom afforded by the device. For example, the audio peak meters described above utilise only a single degree of freedom: the number of lights. To display more information it might be necessary to consider incorporating a further degree of freedom, such as the ability of each individual light to change colour.

Visualising Affect

Much like the audio-level meters discussed above, colour bar representations have been adopted by researchers as a means of displaying affective and social information. Khaorapapong and Purver (2012) featured a colour bar in their design of an 'Icebreaker T-shirt' to be worn at social events (see Fig. 4.2). When two people wearing the t-shirt shook hands, RFID readers on their respective sleeves were used to transfer their online social network details. These details were then compared in order to compute a compatibility score, which was consequently represented on their t-shirts using the colour bar. Hot (red and orange) and cold (green, blue, purple) colours were used to signify high and low compatibility respectively. In this case the colour bar design was chosen as something that would not be too distracting from the social interaction.

Glahn et al. (2009) developed a colour bar visualisation for indicating the amount of activity of individual users in an online community. In another computer-based application, EEG measurements were used to capture valence and arousal aspects of a user's emotions whenever they interacted with a widget within a graphical user interface (Cernea et al., 2013). The detected valence/arousal levels were then mapped onto two colour bars positioned below the corresponding widget. In this case, the segments of the bar represented distinct emotional state readings, and the colours were used to represent the recorded emotion. Red and green were chosen to represent negative and positive valence respectively; and blue to represent high arousal. Following an evaluation of their system, the authors concluded that providing emotional visualisations could enhance window-based interfaces. Cernea et al. (2014) developed a similar system for visualising group affective tone during collaborative interactions on a tabletop interface. They captured levels of valence from EEG headsets and mapped each user's valence level (-1 to 1) to the frequency component of a sine wave. The sine waves for each user were then combined to create a novel visualisation of the degree of group affective tone, which was displayed on the tabletop.

Pietrowicz and Karahalios (2013) describe the development of a real-time tool for visualising expressive aspects of speech during dyadic conversations. Their visualisation comprises a simple graphical display, where the horizontal axis represents time

Figure 4.2: The 'Icebreaker' T-shirt (Khaorapapong and Purver, 2012), designed to unobtrusively signal the compatibility of two people based on comparisons of their social network data. Compatibility level is displayed on the coloured bar at the centre of the t-shirt. Image used with permission.

and the vertical axis represents pitch. Circles are plotted on the graph to represent phonetic features of the spoken words, with the size of the circle representing the sound amplitude. Three phonetic qualities are detected: vowels, obstruent consonants (e.g. s, sh, ch, p, k), and sonorant consonants (e.g. n, m, l, r, w). These are assigned the colours red, blue, and green, respectively.

No existing research was found on the visualisation of performer-derived affective signals during musical performance. The only work encountered was an audiovisual performance involving the creation of a live visualisation based on the heart rate measurements of a chamber orchestra (Votava and Berger, 2012).

Many of the studies discussed above adopt colour as a means of conveying affective meaning. However, more often than not the colour choices appear to be arbitrary, with little consistency across studies. For example, one paper was reviewed where the colour red was used to symbolise high social compatibility; and another where it was used to signify negative valence. Studies that specifically focus on colour and affect refer to 'colour emotion' to describe the emotional feelings elicited by different colours (Lee et al., 2009). Factor analyses have revealed three factors that can be used to parametrise colour emotion space: colour *activity*, colour *weight*, and colour *heat* (Ou et al., 2004; Lee et al., 2009). Ou et al. (2004) provide rankings of 20 colour samples along each of these factors. (Lee et al., 2009) perform factor analysis on emotional aspects of coloured 2D and 3D shapes, from which they identify two additional shape-specific factors: *softness* and *complexity*. The authors also analyse preference, and find that active colours and simple shapes are most preferred.

(Csurka et al., 2010) looked at the associations between colour combinations and emotion. They used existing colour palette databases to extract 15 semantic categories, which were then linked to a range of keywords; many of which have affective associations. In a study of the relationship between affect and the physical properties of colour – hue, chroma, and lightness – Suk and Irtel (2010) asked people to rate their emotional responses to colours, using the self-assessment manikin (SAM). The SAM is a commonly used method for collecting subjective ratings of the three dimensions of affect: arousal, valence, and dominance. Each dimension is represented using a series of images of a simple graphic character (SAM). For example, the valence dimension shows characters ranging from smiling to frowning. The authors found that chroma was positively correlated with all three affective dimensions. Additionally, bright colours were found to score highest for valence and dominance, whilst blue was rated highest for dominance and valence, but lowest for arousal. The study also looked at the difference between presentation medium, using a CRT monitor in one experiment, and printed sheets in the other. No systematic effect of media was observed.

Physiological responses to colour have been investigated, but no consistent relationship between the two has been found (Kaiser, 1984; Detenber et al., 2000; Suk and Irtel, 2010).

4.2.5 Summary and Relevant Findings

To summarise this section, key findings are discussed in relation to each of the four topics that have been reviewed: gaze, motion, cardiac activity, and visual feedback.

Gaze

Gaze can be measured using eye-tracking headsets or desktop devices. The former offer the highest accuracy and freedom of movement, and in recent years affordable devices have been developed. Commonly used eye-tracking features are the number of fixations (looking at a fixed point); and the duration of time spent looking at a particular area of interest, expressed either as a percentage of the total time, or as a mean value.

Studies have highlighted the important functions of gaze in non-verbal communication (NVC), but have predominantly investigated this in the context of conversational interactions. There is an absence of empirical work on the use of gaze in collaborative music making. With regard to applications for eye-tracking in collaborative interactions, existing work has focused on supporting screen-based collaborative work, with mixed results. In the context of musical interactions eye-tracking has been used as a control input for musical interfaces. There are no known studies involving the use of eye-tracking for supporting co-present musical interaction.

Motion

Motion can be captured using marker-based, electromechanical/electromagnetic, vision-based, and accelerometer-based methods. Whilst the first two methods offer the highest accuracy, the latter two options are better suited to applications where low cost, mobility, and minimal set up time are prerequisites. Kinematic motion features can be extracted as low-level time-derivative features (e.g. velocity, acceleration), or high-level time series features (e.g. dynamic motion patterns). Quantity of motion (QoM) has been shown to be a useful feature for discriminating between high and low arousal emotions, and for recognising engagement during interactions.

During musical performance, movement can serve numerous functions, such as producing sound, or communicating non-verbal signals. This thesis is predominantly concerned with the latter, due to its relevance to affective and behavioural expression. Existing studies have used motion capture to analyse the nature of non-verbal interaction between collaborating musicians; and to provide musicians with performance-related feedback. However, the use of motion capture for providing affective or behavioural feedback to collaborating musicians has not been explored.

Cardiac Activity

Accurate measurements of cardiac activity can be obtained using ECG sensors. However, more recent advances have seen heart rate (HR) sensors integrated onto wearable devices, such as smartwatches. HR and heart rate variability (HRV) are commonly used cardiac activity features. With respect to the latter, additional time and frequency-domain features can be extracted to analyse cardiac functioning in detail.

Existing work on cardiac activity was reviewed in relation to creativity, decision making, and musical performance. Literature on creativity suggests that contrasting thought processes are involved in creative cognition, and it is feasible that these thought processes could be accompanied by distinct phases of cardiac activity. There is a lack of existing research on this topic; however, studies have evidenced ways in which cardiac activity relates to more general decision making processes.

Cardiac activity has been investigated in the context of musical performance, where it has been shown to relate to self-reported flow; and to vary according to the difficulty of the music and the expressivity of the performance. When listening to music, cardiac features have also been shown to change in relation to tempo and emotional content.

Visual Feedback

Work was reviewed relating to the provision of minimal, dynamic, visual feedback. A common representation of such information in the musical domain is the audio level display; consisting of a strip of coloured lights that represent the instantaneous volume level. Colour bar representations have also been adopted by researchers as a means of communicating affective and social information. However, this has not been explored in the context of musical performance.

Studies have provided evidence for consistent associations between colours and emotion. In particular, chroma was shown to be positively correlated with the arousal, valence, and dominance dimensions of affect. However, the use of colour in HCI appears to be somewhat arbitrary, and based upon subjective preference, rather than empirically-derived rules.

In summary, this section has reviewed topics that concern the input measures and output modalities that were identified at the beginning of this chapter. The following section introduces the **LuminUs**: a device that is designed to use affective and behavioural sensing to provide collaborating musicians with real-time visual feedback. The design of this device is informed by the work reviewed in this section.

4.3 The LuminUs: Overview

This next two sections in this chapter describe the design of a device to support the affective and social aspects of collaborative music making. The device was named the **'LuminUs'**, since its intention is to *illuminate* the 'us' in collaborative music making. The design of the device is based upon the idea of providing each musician with a visual representation of their collaborator's affective state and social signals. This section provides an overview of the device, including justifications for design choices.

4.3.1 Sensors and Inputs

The previous section included an in-depth review of measurement techniques, features, and relevant research relating to three affective measures: gaze, motion, and cardiac activity. All three of these measures are incorporated into the design of the LuminUs. However, at this stage measures of cardiac activity are not used to provide real-time feedback. Findings from Study 1 (see Chapter 3) suggested that cardiac activity may relate to the musical decision making processes of performing musicians. However, in order to provide meaningful feedback, more understanding is required of the nature of these relationships. Incorporating cardiac activity as a latent measure facilitates further investigation into ways that it could be used to provide feedback to collaborating musicians in the future.

Specific sensor devices were chosen for the collection of gaze, motion and cardiac data. This section briefly describes these devices, and the justifications for choosing them:

Gaze sensor: The Pupil[2] eye-tracking headset (see Fig. 4.3(a)) and software (Kassner et al., 2014) were chosen to collect real-time gaze data from musicians. The device has been developed over the last four years as an open-source hardware and software project. It tracks the movement of the right eye using a single infra red camera, and simultaneously captures the wearer's field of vision (FOV) using a separate forward-facing video camera. Following a short calibration (~20 s), the software is able to map the gaze point of the wearer onto the live FOV video. One of the advantages of the Pupil device, compared to other commercial eye-tracking glasses, is that it is inexpensive (at the time of writing, roughly 3% of the price[3]). This meant that multiple devices could be purchased for use in experiments with two or three musicians. Furthermore, the device has some distinctive

[2] http://pupil-labs.com
[3] Compared to the Tobii Glasses 2 - http://www.tobii.com

(a)

(b)

Figure 4.3: Images of the devices used to collect data from collaborating musicians - (a) the Pupil eye-tracking headset; and (b) Shimmer ECG/Motion Sensor

practical benefits compared to the alternatives. At 50 grams, it is the lightest eye-tracking headset currently available. It also uses open source software, enabling the development and incorporation of features that are application-specific. One disadvantage to this device is that it is not currently wireless, meaning that the musicians are somewhat restricted in their movement. However, the release of affordable wireless eye-tracking headsets is foreseen in the near future.

Motion and cardiac sensor: The same device that was used in Study 1 was selected for collecting body motion and ECG data. The Shimmer[4] sensor incorporates a tri-axial accelerometer, and collects ECG measurements using four electrodes attached to the user's chest. From a practical standpoint, the inclusion of two sensors in a single device simplifies data collection. Furthermore, the Shimmer device is wireless, which means that the data can be accessed in real-time through a Bluetooth connection.

4.3.2 Signal Processing

Informed by the work reviewed in the previous section, the decision was made to extract the following gaze and motion features in order to provide real-time feedback:

Eye-tracking feature: The Pupil eye-tracking software has the ability to recognise fiducial markers using the FOV camera. It can then calculate the coordinates of the wearer's gaze fixation, relative to the detected marker. This makes it possible to place markers on the headsets of each musician, and consequently register the

[4]http://www.shimmersensing.com

moments when one person is glancing at the other. A threshold can be used in order to specify how close the gaze point must be to the marker in order for a glance to be registered. This results in a binary feature: glancing, or not glancing.

Motion feature: Similar to the first study, the decision was made to use an overall measure of the quantity of motion (QoM) of the musician's body. This consists of the difference between consecutive summations of the absolute X, Y, and Z accelerometer values. Smoothing is applied to the resulting signal in order to give an averaged measure of the musician's body motion.

4.3.3 Feedback Visualisation

A colour-bar method was selected for displaying feedback to the musicians. The colour-bar is commonly used for displaying sound levels in audio and is likely to be a representation of dynamic information that musicians are familiar with. Simplicity and familiarity were important factors, since one of the aims was to avoid introducing unnecessary clutter or distraction to the collaborative music making environment. This was a problem cited by numerous authors who attempted to provide eye-tracking feedback during collaborative group work (see Section 4.2.1). Furthermore, the review of existing research in Section 4.2.4 indicated that the colour-bar is one of the most popular methods of conveying affective information during interactive situations. An early illustration of the LuminUs is shown in Fig. 4.4. It consists of a single strip of addressable RGB (red, green, blue) light emitting diodes (LEDs). This means that the colour and illumination of each LED can be controlled individually, which provides flexibility for testing different designs.

Regarding the mapping of input signals to the colour-bar, the decision was made not to attempt to represent gaze and motion information simultaneously, as this could be confusing for the musician. Instead, the device was developed such that the type of feedback (gaze or motion) could be selected using the software. In the gaze feedback mode the colour-bar is used to indicate to the musician when their collaborator is glancing at them. Specifically, the number of illuminated lights increases as the length of the glance increases. In the motion feedback mode the quantity of motion (QoM) value is simply mapped to the number of illuminated lights, such that the more the musician is moving, the more lights are illuminated.

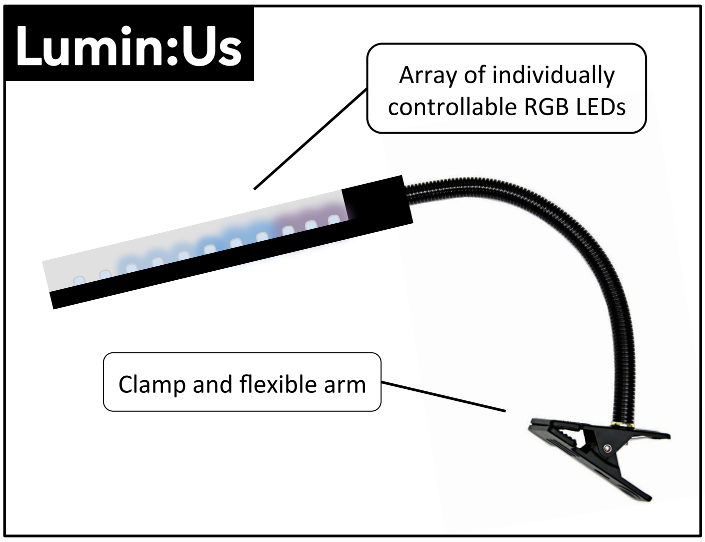


Figure 4.4: Early design sketch for the LuminUs - a light for illuminating affective and behavioural features of collaborative music making.

4.4 The LuminUs: Hardware and Software

This section provides a detailed description of the design of the LuminUs. It is separated into a description of the hardware components, followed by the software used to operate the LuminUs.

4.4.1 Hardware

Lighting

The main requirements for the lighting were to have a set of individually controllable RGB LEDs, which were small enough to be positioned close together in a single row. A further requirement was that the intensity of the light output could be adjusted. The NeoPixel stick[5] (see Fig. 4.5) was chosen, which consists of eight small (5 mm × 5 mm) RGB LEDs mounted on a single strip of printed circuit board (PCB). The NeoPixel LEDs are individually addressable, since each has its own driver chip integrated into the LED. This means that all eight LEDs can be controlled using a single microcontroller signal. The NeoPixel sticks can also be linked together to create longer LED strips. The decision was made to use a pair of sticks for the LuminUs, giving a total of 16 LEDs.

[5] http://proto-pic.co.uk/neopixel-stick-8-x-ws2812-5050-rgb-led-with-integrated-drivers/

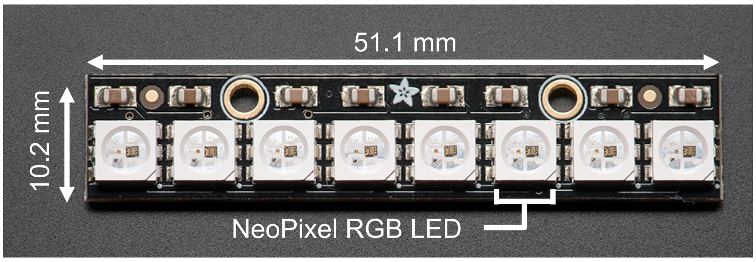


Figure 4.5: The NeoPixel stick, used for the light output of the LuminUs.

Microcontroller

In order to control the RGB LEDs, a compact programmable microcontroller was required, with serial connectivity, the ability to output a bitstream at 800 kHz, and a regulated 5V DC power output. The Arduino-compatible Pro Micro[6] microcontroller from SparkFun was chosen (see Fig. 4.6). The device features an ATmega32U4 microcontroller that runs at 5V/16MHz. and has 12 digital input/output pins. It has USB connectivity, which means that it can be programmed and powered from a computer. This also allows computer programs to communicate with the microcontroller via serial port. In the case of the LuminUs, this means the light output can be controlled directly from a computer, based upon the gaze and motion data.

The NeoPixel sticks and Pro Micro were connected as shown in Fig. 4.7. The three necessary connections are the 5V DC and ground (GND) connections to provide power, and the data input (DIN) to control each individually addressable LED. In turn, the USB connection to the Pro Micro provides power to the microcontroller and allows the LEDs to be controlled from a computer.

[6]http://proto-pic.co.uk/pro-micro-5v-16mhz/

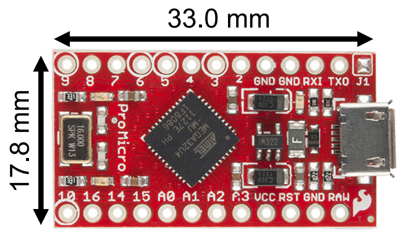


Figure 4.6: The Pro Micro microcontoller, used to communicate with and control the lights on the LuminUs.

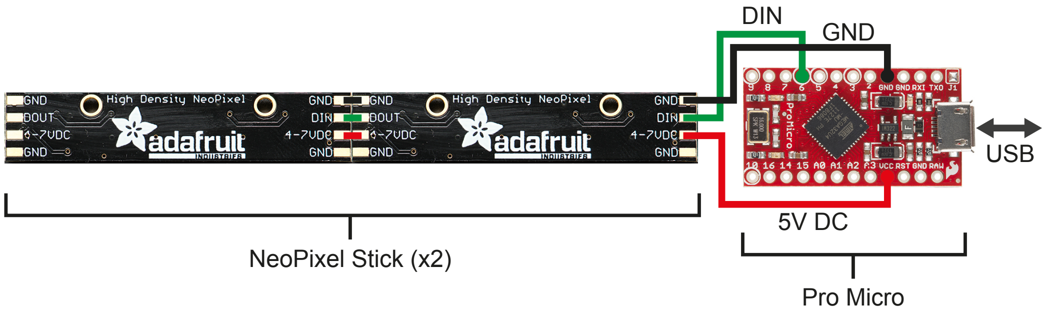


Figure 4.7: Diagram showing the connections between the two NeoPixel LED sticks and the Pro Micro microcontroller.

Casing

A custom case was designed and built to hold and protect the LuminUs circuitry, and to enable the device to be mounted on a flexible stand. The casing was built from six layers of laser cut acrylic, and its dimensions are 145 mm × 30 mm × 26 mm. An exploded view of the casing design is shown in Fig. 4.9. The case is held together with six bolts, and a gap in the casing provides access to the USB port on the microcontroller. The two central layers have a small protrusion with a hole in the centre, enabling the casing to be connected to a flexible arm. A frosted, transparent perspex window was used to diffuse the light from the LEDs, making it less intense to look at.

Two mirrored versions of the LuminUs casing were built. This was done so that, when two musicians stand side by side, each of their LuminUs devices can be oriented such that it points towards their partner. Views of the assembled casing are provided in Fig. 4.8.

4.4.2 Software

The software required to make the LuminUs functional comprises five separate programs: i) Shimmer; ii) Pupil eye-tracking; iii) LuminUs control; iv) LuminUs server; and v) microcontroller software. Each of these programs has distinct roles and data are passed between them via network communication. Figure 4.14 provides an overview of the five programs which are required to operate the LuminUs with either gaze or motion feedback. Some of these programs were developed specifically for use with the LuminUs, whilst others are modified versions of existing software. Detailed descriptions of each of the programs are provided on the following pages.

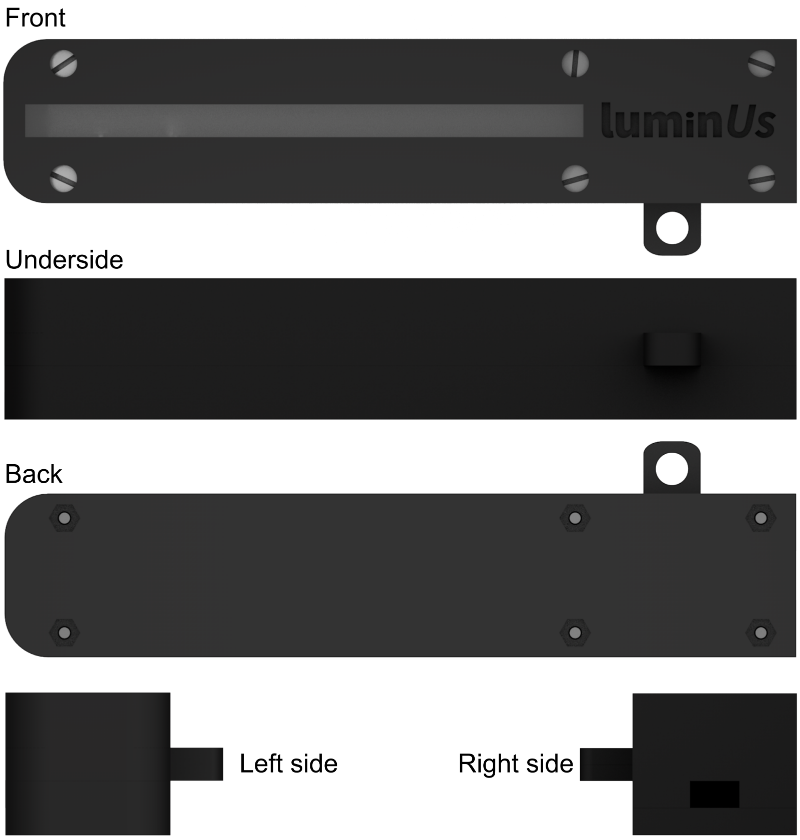


Figure 4.8: Model views of the assembled LuminUs casing.

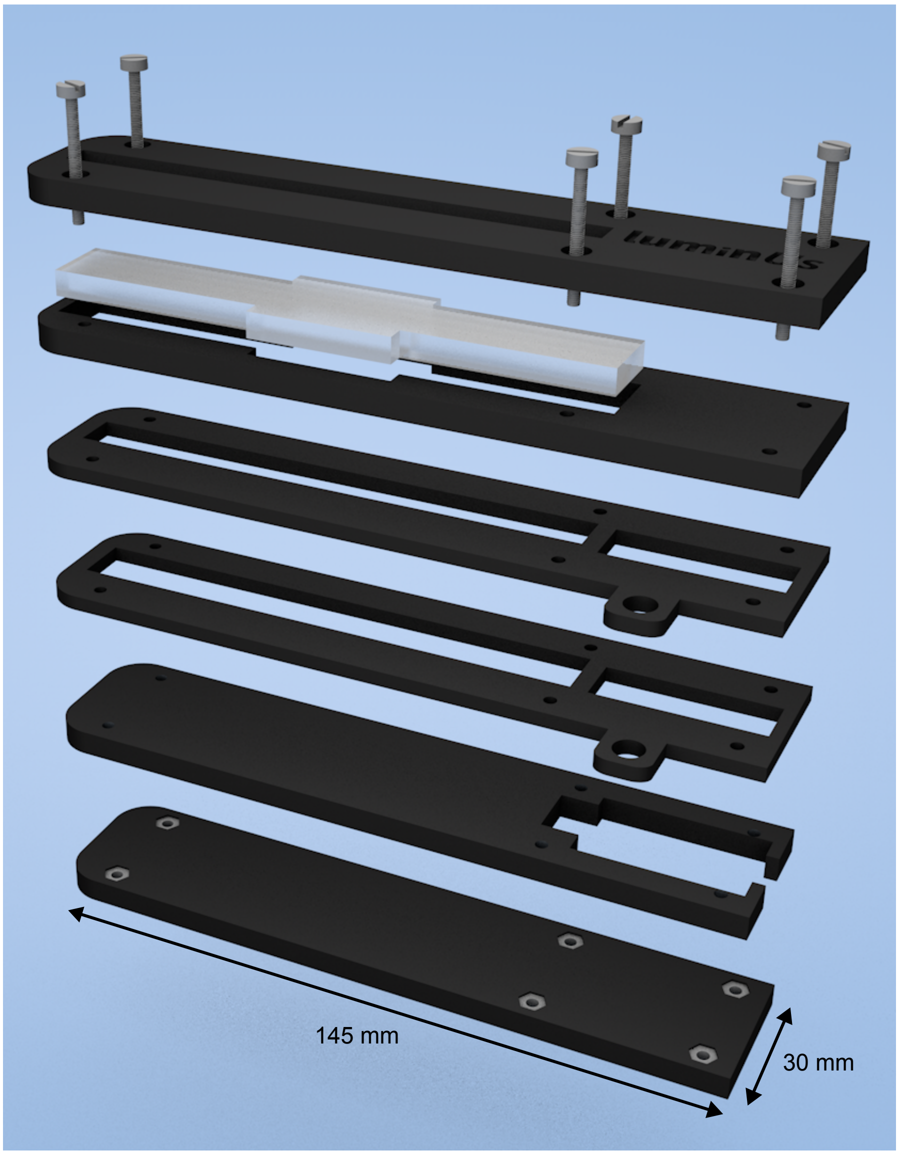


Figure 4.9: Exploded model of the LuminUs casing, consisting of four layers of 5 mm acrylic with two 3 mm layers at the centre. A transparent window of clear acrylic was used to diffuse the light from the LuminUs.

Shimmer Software

The Shimmer software was written in Matlab using the API provided by Shimmer. The software graphical user interface (GUI) (see Fig. 4.10) allows the user to specify an IP address to transmit the data to; and to provide the wireless COM port numbers for two Shimmer devices. Once the user clicks 'START', the software attempts to establish a connection with the Shimmer sensors. Messages appear in the 'Progress Messages' box to inform the user whether the connections have been successful. If successful connections are made then the software sends initial Open Sound Control (OSC) messages containing the current time-stamp on each Shimmer device. The time-stamp refers to the current time in milliseconds on the internal hardware clock of the Shimmer sensor. This time is used as a reference time by the Max MSP software, in order to synchronise the Shimmer data with other data sources. An additional button on the GUI, labelled 'Send Time Sync', enables the user to manually send new time-stamps from the devices.

Once the initial time-stamp data have been sent, the software continuously attempts to grab new data packets from each of the Shimmer devices. These data packets contain the calibrated data from the ECG and accelerometer sensors. Once a packet is received, the software sums the absolute values of three accelerometer axes (X, Y, and Z), and then sends these data along with the ECG data in a single OSC message. The first part of this OSC message contains a tag, which identifies which of the two Shimmer devices the data originates from: either ECG1 or ECG2. An example of a single data packet is given below:

$$\underbrace{\text{/ECG1}}_{\text{Tag}} \quad \underbrace{15272.732}_{\text{Time-stamp}} \quad \underbrace{-0.063}_{\text{ECG lead 1}} \quad \underbrace{-0.075}_{\text{ECG lead 2}} \quad \underbrace{17.803}_{|X|+|Y|+|Z|}$$

Pupil Eye-tracking Software

The Pupil eye-tracking platform includes a fully developed open-source software package (see Fig. 4.11). The software receives video streams (via USB connections) from the FOV and eye cameras on the Pupil headset and has the following features:

Visualisation: The video streams from both of the headset cameras can be visualised. This allows the user to check that the eye-camera is correctly positioned to detect the wearer's pupil; and that the tilt of the FOV camera is representative of the wearer's field of view. Parameters, such as the focus and exposure, can be manually adjusted for both cameras.

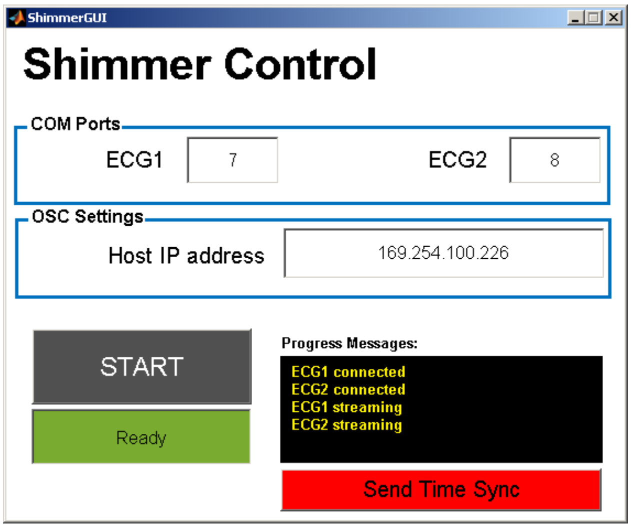


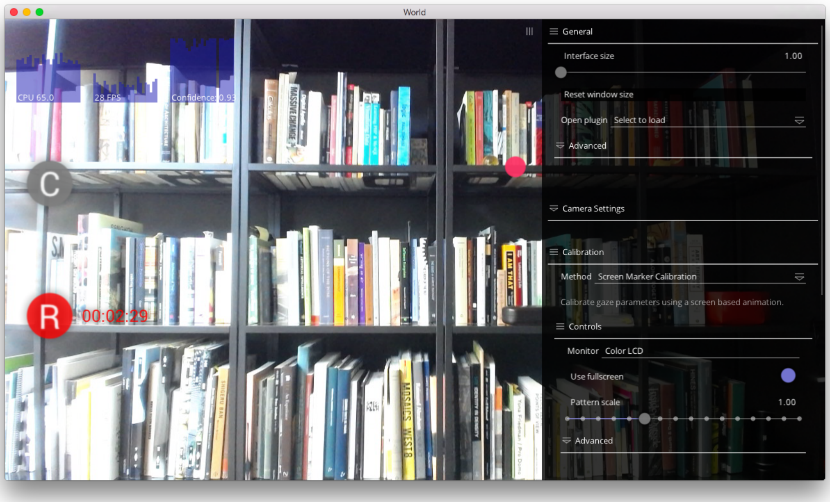
Figure 4.10: The software graphical user interface (GUI) created for the Shimmer sensors.

Checking pupil detection: The user can check that the pupil is being correctly detected (see Fig. 4.11(b)) and can adjust various parameters, such as the maximum and minimum pupil size, in order to optimise the detection performance.

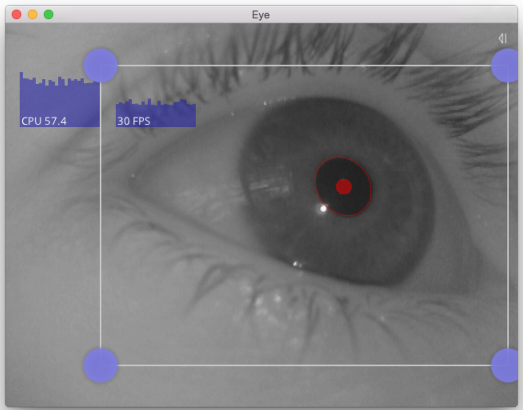
Calibration: In order to correctly identify where the user is looking, the pupil software must correlate the detected pupil positions with locations in the field of view. A short screen marker calibration can be performed, which requires the user to follow a small marker on the screen with their eyes, whilst keeping their head stationary. Once the calibration is complete the user's current gaze potion will appear as a red dot overlaid on the FOV camera stream (see Fig. 4.11(a)).

Recording: The FOV video stream and corresponding gaze and pupil position data can be recorded and saved using the software.

Additional functionality had to be added to the Pupil software to enable it to detect when a user is looking at a marker in their FOV; and communicate this with the LuminUs control software. The marker detection was achieved using existing code in the Pupil software that performs the detection of the circular calibration marker (see Fig. 4.12). This code was modified to create a plug-in that detects the marker and calculates a value representing the normalised distance between the user's current gaze position and the position of the detected marker. Existing Pupil networking plug-ins were then modified to enable this distance value to be sent to the LuminUs control software via an OSC message.



(a)



(b)

Figure 4.11: Images of the Pupil software interface: (a) main window with POV camera view; and (b) eye-camera window.

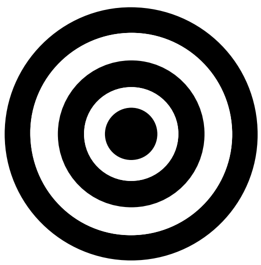


Figure 4.12: The circular marker used for gaze detection.

LuminUs Control Software

The functions of the LuminUs control software are to receive ECG, accelerometer, and gaze data; process the accelerometer and gaze data; determine the appropriate light output for the LuminUs; and send the lighting control data to the LuminUs server. To achieve this, custom software was developed using the Max MSP[7] visual programming language. The control software receives data packets from the Shimmer software and Pupil software via OSC protocol. ECG data are simply saved to disk (for post-hoc analysis). The accelerometer and gaze data are processed as follows:

Accelerometer data processing: Each accelerometer value is taken from its Shimmer data packet and down-sampled. The difference between consecutive down-sampled accelerometer values is then calculated to provide an overall rate of change of acceleration (jerk). This step is performed in order to remove the effects of gravitational acceleration. The jerk values are subsequently smoothed using a logarithmic filter. Each smoothed value is then mapped to the range 0-16, which represents the number of LEDs on the LuminUs that should be lit up (0 means no LEDs are lit).

Pupil data processing: Each gaze distance value provided by the Pupil software is compared to a distance threshold value. If the value is less than the threshold then a timer is started, which counts up towards a maximum value of 2000 ms. This value was based upon the mean glance length from Study 1, which was 2100 ms. The timer value is then mapped to an inverse exponential curve, which has an output range between 0 and 16; again, representing the number of LEDs that should be lit on the LuminUs. This is done so that the number of lit LEDs increases quickly when a gaze is detected, and then becomes more gradual as the length of the gaze increases. If at any point the gaze distance value rises above the threshold then the timer is set to count downwards towards zero.

The output from the Pupil and Shimmer data processing consists of two numbers in the range of 0-16, representing the number of LEDs that should be lit for gaze and motion feedback respectively. The motion feedback value is shifted by 17, so that it falls within the range of 17- 33, whilst the gaze value remains within the range of 0-16. The software then allows the user to select which value is sent to the LuminUs, and consequently, which type of feedback is provided. This control value is transmitted as a

[7] https://cycling74.com/

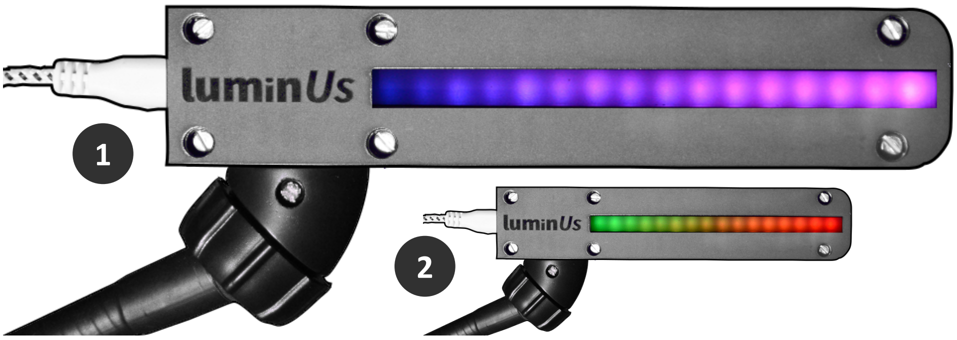


Figure 4.13: The LuminUs with all LEDs lit in gaze feedback mode (1) and motion feedback mode (2).

UDP message to a given IP address. Data packets are processed continuously, resulting in a stream of control values.

LuminUs Server Software

This is a simple program which receives UDP messages from the LuminUs control software and forwards them on to the LuminUs device as serial messages. This program must be run on the same computer that the LuminUs device is connected to. Separating the server software from the control software means that the LuminUs can be connected to one computer and controlled from another computer on the same local network.

Microcontroller Software

This software runs on the Pro-Micro microcontroller inside the LuminUs device. It was written in C++, using the Arduino Intergrated Development Environment[8]. The program receives serial data from the LuminUs server software, which are simply values in the range of 0-33. If the values are in the range 0-16 then the software treats them as gaze feedback data; otherwise they are treated as motion feedback data. The values are then used to determine how many LEDs should be lit, and what colour the LEDs should be. For the gaze feedback the lights were programmed to gradually increase from violet to pink. For the motion feedback they were programmed to increase from green to red. Figure 4.13 shows the state of the lights for each type of feedback when all the LEDs were lit.

[8]https://www.arduino.cc/en/Guide/Environment

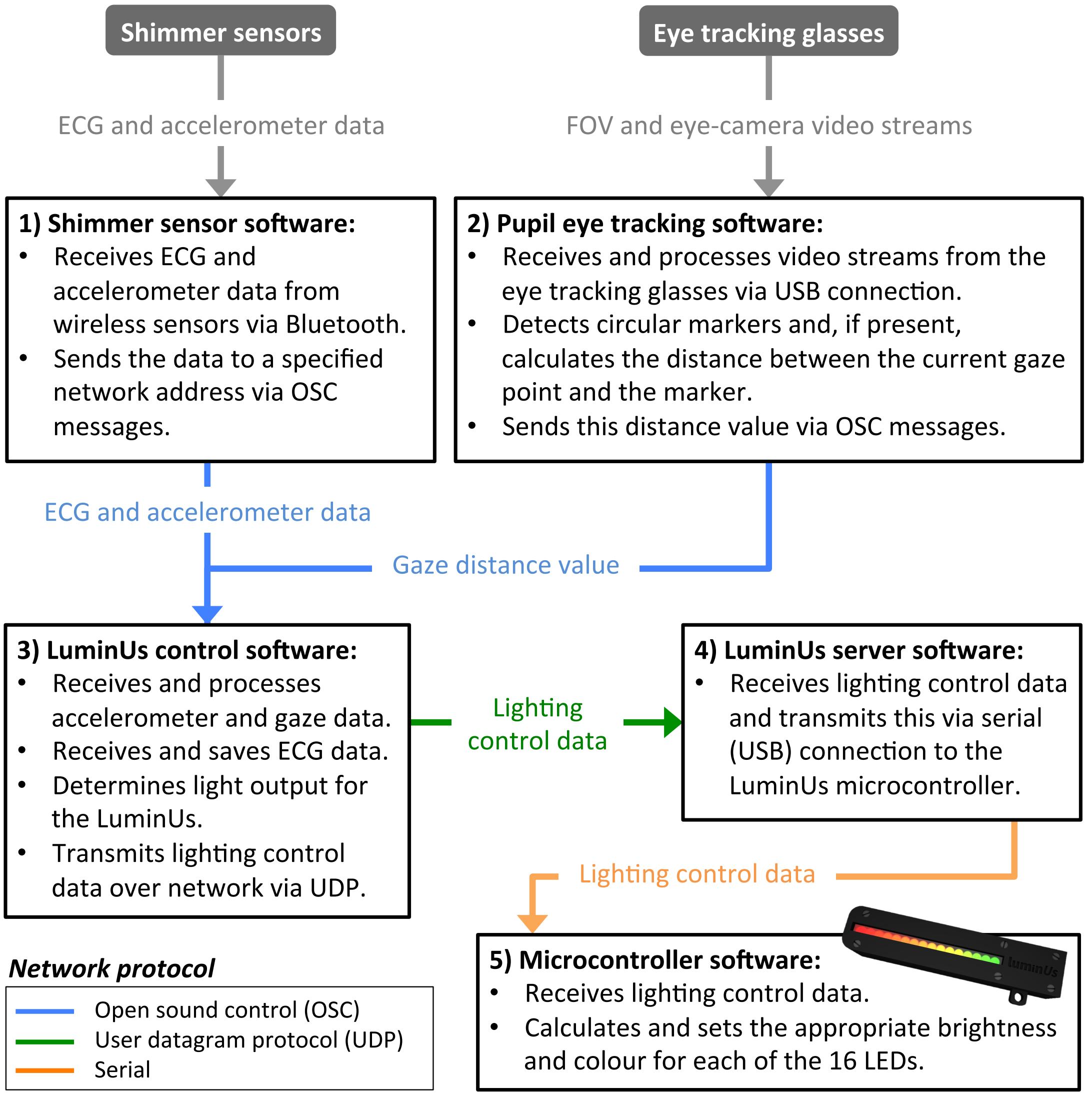


Figure 4.14: A system diagram, showing the five software programs that are used to operate the LuminUs.

4.5 Discussion

The LuminUs is a prototype device, which was inspired by the findings reported in Chapter 3. In the present chapter, these findings informed a comprehensive review of gaze, motion, and cardiac measures; including methods for their measurement, associated features, and relevant existing research. Consequently, the selection of the sensors and features, on which the LuminUs is based, is supported by an in-depth understanding of related work and issues. With regard to the provision of affective and behavioural feedback, the approach taken in the design of the LuminUs is to provide the musicians with a simplistic visual representation of the extracted features. In this sense, the device does not attempt to identify and represent *specific* affective and behavioural states. Instead, it is designed to provide sensor-derived information, which is open to interpretation by the beholder. In affective computing research, a common approach is to design systems that classify continuous affective inputs according to specific labels, categories or dimensions (Gunes and Schuller, 2013). One of the limitations of this approach is that affective and behavioural signals may be highly context dependent. For example, increased movement could indicate agitation in one context, or elation in another. By leaving some aspects of the interpretation open to the user, the LuminUs is less susceptible to these limitations.

The design requirements set out at the start of this chapter (see section 4.1.2) specified that the feedback provided by the device should be visual, minimal, and dynamic. The review of visual feedback methods in section 4.2.4 indicates a lack of existing research to inform the design of visual feedback devices for performing musicians. The simple, light-based design of the LuminUs offers the flexibility to modify parameters such as colour and brightness; and to create varying patterns of dynamic feedback using combinations of lights. This makes the device suitable for testing and evaluating various feedback configurations.

The LuminUs design has a number of limitations. In particular, it requires multiple software programs and, subsequently, computers in order to collect and process the sensor data. This means that, whilst the device is small and portable, the entire system is not. Additionally, the device is designed with dyadic interactions in mind, which means that it is limited to providing affective and behavioural feedback from a single musician.

Finally, the work reviewed in this chapter, and the exploratory findings in Chapter 3, indicate that interesting relationships could exist between cardiac activity and musical decision making. Whilst the LuminUs is able to collect ECG data, it does not use these data to provide real-time feedback. This is due to the fact that there is currently insufficient evidence to inform the selection of specific features and mappings that might provide collaborating musicians with meaningful feedback. Instead, cardiac activity is incorporated into the LuminUs as a latent measure; thus facilitating further investigations of cardiac activity in the context of collaborative music making.

4.6 Summary

This chapter established specific considerations for the design of a sensor-based device to support and enhance collaborative music making. It reported a comprehensive review of relevant methods and research; which led to the design and development of a prototype device - the LuminUs. In the following chapters, two further studies are reported. Study 2 investigates the effects of the LuminUs feedback upon collaborative music making, and tests specific hypotheses concerning relationships between cardiac activity and decision making. This study uses controlled settings and short-term interactions. Study 3 evaluates qualitative aspects of musicians' use of the LuminUs over the course of a longitudinal study.

Chapter 5

Study 2

Testing the LuminUs in the Lab

Informed by the findings from Study 1 and the subsequent review of related work in Section 4.2, the previous chapter documented the design and development of the LuminUs: a device for providing collaborating musicians with real-time visual feedback about the gaze or body motions of their co-performers. Gaze and body motion were chosen because they both serve important functions in the expression and communication of non-verbal social and affective signals. The LuminUs collects gaze data from eye-tracking headsets and visual markers worn by the musicians. These data are then processed and mapped, such that the lights on a musician's LuminUs will illuminate whenever their co-performer looks at them. Motion data are collected from accelerometers and processed such that the number of illuminated lights on a musician's LuminUs provides a representation of how much their co-performer is moving.

The LuminUs also incorporates an ECG sensor, enabling cardiac activity to be recorded as a latent measure (i.e. one that does not influence the visual feedback). The inclusion of this measure was inspired by findings from Study 1 and reviews of related work, which suggest that cardiac activity could be used as an indicator of musical decision making; providing musicians with an enhanced awareness of their collaborator's intentions. However, further work was deemed necessary before this measure could be meaningfully mapped to the visual feedback of the LuminUs.

The present chapter reports a study in which the LuminUs was put to the test with collaborating musicians in a controlled environment. The purpose of this study was to evaluate the influence of the existing LuminUs feedback modalities upon collaborative music making; and to consider the potential for cardiac activity-related measures to be used in future versions of the LuminUs.

5.1 Aims

In contrast to the first study, which had an exploratory design, the present study was designed to investigate specific questions and hypotheses relating to two aims:

1. To investigate the **effects of the LuminUs feedback** upon aspects of co-present interactions between musicians.

2. To undertake further investigation of potential relationships between the **cardiac activity and musical decision making** processes of musicians.

With respect to the first aim, exploratory questions were posed relating to the effects of the LuminUs feedback. These are outlined in the following section. Section 5.1.2 then outlines specific hypotheses with respect to the second aim.

5.1.1 On the Effects of the LuminUs (LQ)

Given the lack of previous research on the provision of real-time affective and behavioural feedback to performing musicians, specific exploratory questions were devised. Concerning *gaze or motion feedback* to the following questions were posed:

LQ1 Does the LuminUs influence the musical outcomes of the collaborations?

LQ2 Does the LuminUs influence self-reported measures of the musicians' enjoyment and interaction with their co-performer?

Furthermore, concerning *gaze feedback* to the following questions were posed:

LQ3 Does the gaze feedback influence the number of glances exchanged between musicians during collaborative interaction?

LQ4 By facilitating communication between musicians, does the gaze feedback influence the number of musical changes made during a composition?

Finally, the following questions were posed concerning the provision of *motion feedback*:

LQ5 Does the motion feedback have an impact upon the overall amount of motion during collaborative interaction?

LQ6 Does the motion feedback stimulate increased awareness of the other participant, as indicated through more glancing at the other?

5.1.2 On Cardiac Activity and Musical Decision Making (cH)

Given the findings from Study 1 and the existing research discussed in the previous chapter, the following hypotheses were devised concerning relationships between cardiac activity and musical decision making:

cH1 Levels of the LF/HF ratio of heart rate variability will be correlated with the number of musical decisions. LF/HF ratio has been shown to relate positively to self-reported flow (de Manzano et al., 2010; Gaggioli et al., 2013), as well as levels of expression in musical performance (Nakahara et al., 2009). Since the number of musical decisions during a performance is likely to be correlated with both flow and expressiveness, LF/HF ratio will be greater during periods where more musical changes have taken place.

cH2 Heart rate variability will be negatively correlated with the number of musical changes. Numerous studies have reported that HRV decreases in response to high mental workload and cognitive focus (Hjortskov et al., 2004). These cognitive functions are expected to increase when musical changes are made, therefore leading to a negative correlation between HRV levels and measures of musical change.

cH3 Musical decisions will tend to coincide with heart rate extrema. The findings from Study 1 indicated that for some participants there was a significant relationship between rhythmic change points and heart rate extrema. It is hypothesised that this relationship occurs because people either introduce change when they are bored (low arousal/heart rate), or when the part they are playing is too challenging (high arousal/heart rate).

5.2 Research Design and Data Collection

In order to test the questions and hypotheses outlined in the previous section, a controlled experiment was designed, where pairs of musicians were asked to create short improvised accompaniments to an animation under different conditions. In each condition the LuminUs would either provide motion feedback, gaze feedback, or no feedback at all. The research design shared a number of similarities with the first study. Paired musicians (dyads) were studied, employing a within-subjects design, where data were collected from each dyad in each experimental condition. The experiment was also designed to be as close to a real-world interaction as possible, whilst still enabling the collection of controlled measures. In contrast to the first study, the pairs of musicians consisted of a percussionist and a pianist. This was done to extend the investigations from minimalist, single-instrument interactions, to more complex and conventional, mixed-instrument interactions. This decision was facilitated by the fact that there was less desire to control for specific instrument-related aspects of the experiment, such a physical exertion.

Another major difference to the first study was the introduction of a visual stimulus to guide and influence the musicians' improvisations. Dyads were asked to create musical accompaniments to a two minute long animation. Details of the selection and use of this animation are provided in this section; along with a thorough description of the research design and data collection methods.

5.2.1 Animation

The main experimental task in this study was to create a live, improvised accompaniment to a silent animation. A visual stimulus was selected for the following reasons: i) to introduce some control over the improvisations, such that the musicians might be influenced to change their playing at similar points because of distinct visual cues in the animation; ii) to encourage the musicians to work collaboratively, heightening the need for mutual awareness and non-verbal communication; iii) to increase the external stimuli beyond that of the instrument and co-performer, as would often be the case in real-world performances. An animation was chosen to avoid images of human subjects, which could evoke more substantial and varied affective responses.

Four animations were short-listed for use in the study (see Table 5.1 and Fig. 5.1). These were selected from hundreds of animations on the video sharing website, Vimeo[1]. The main criteria for selection were that the animation had *visual continuity* (to avoid sudden visual changes); that it included some distinct *theme changes* (to stimulate musical change); and that its content *was not too emotive.* Permissions to use each animation for research purposes were obtained from the copyright holders.

[1]https://vimeo.com/

Table 5.1: Details of the four short-listed animations.

	Title	**Creator**	**Link**
A	Float	Haruki Kawanaka	vimeo.com/78550258
B	Spherikal	Ion Lucin	vimeo.com/39792837
C	Killing the Work	Juliette Hamon-Damourette	vimeo.com/20247426
D	Accumulonimbus	Andy Kennedy	vimeo.com/13846037

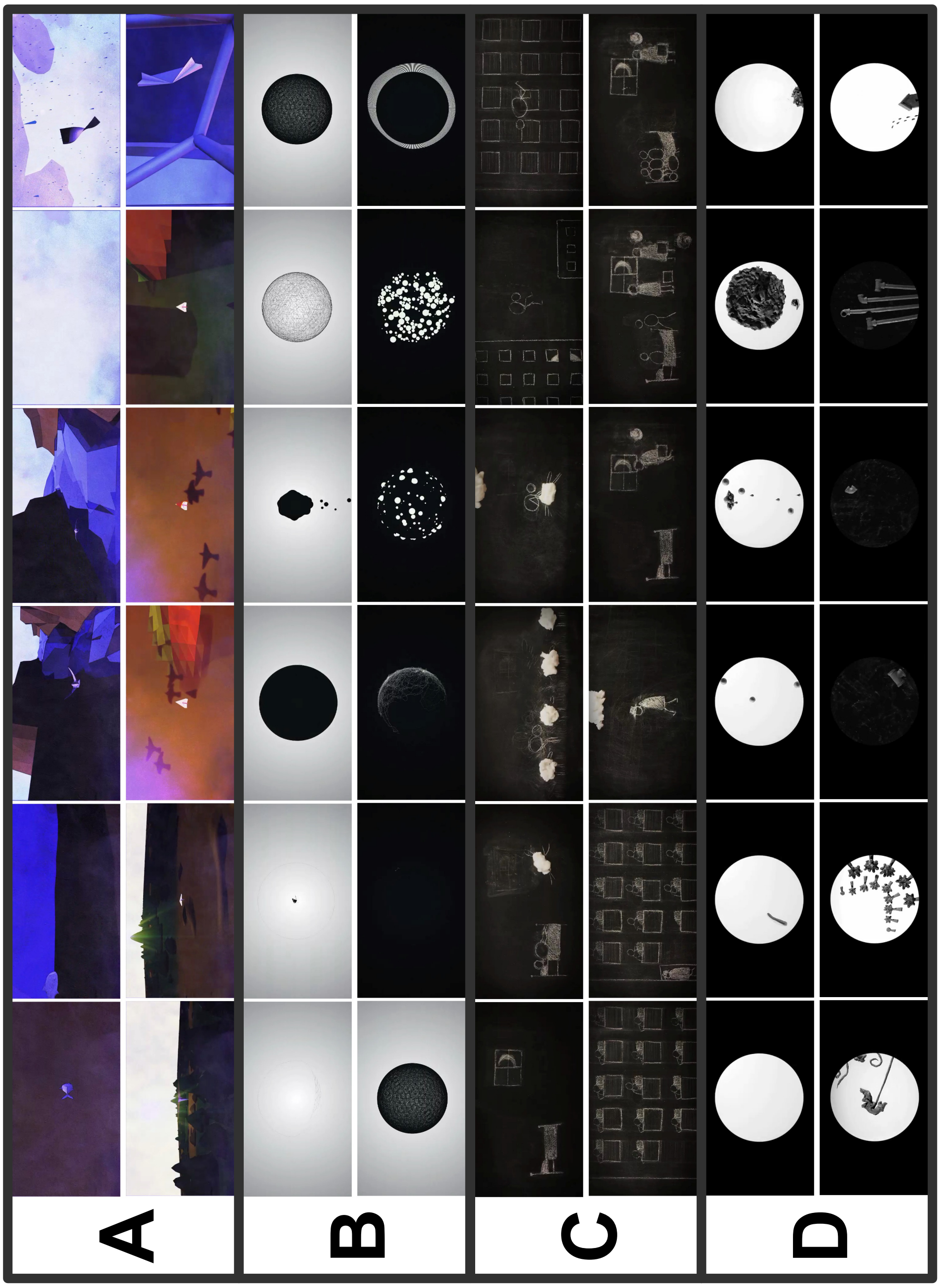


Figure 5.1: Film strips of the four short-listed animations. Frames were captured every 10 seconds. See Table 5.1 for details on each animation.

A short survey was carried out in order to select which animation to use in the study. Musicians were invited to watch and rate the animations based upon their suitability for improvised musical accompaniment. Each animation in the survey was cut down to two minutes in length and the animations were presented in random order, without sound. Responses were collected from 24 musicians, none of whom were participants in the present study. Respondents had an average musical experience of 17 years. The results are shown in Table 5.2. The animation entitled 'Float' (animation A) received the highest or joint-highest score for each question and was consequently selected for use in this study.

Table 5.2: Animation survey responses.

	Animation			
Question	A	B	C	D
1. In your opinion, how well suited is this animation for improvised musical accompaniment?	**4**	3	3.5	3
2. How musically inspiring did you find this animation?	4	3	4	4
3. How visually engaging would you find this animation if you were improvising to it with another person?	4	4	4	3
4. Please rank the four animations in order of preference.	**2**	2.5	3	3
Note: Scores for questions 1-3 are median 5-point Likert scale ratings. The score for question 4 is the median ranking (1 = highest). Total responses = 24.				

5.2.2 Participants

15 percussionists and 15 pianists were recruited from music colleges in London via email advertisements. The participants comprised 7 females and 23 males, aged between 18 and 38 (M = 22.7, SD = 4.7). Their playing experience ranged from 3 to 33 years (M = 11.6, SD = 5.6). Participants were assigned to percussionist-pianist pairs, such that the individuals in each pair did not know one another. This was done to avoid introducing non-random variables pertaining to existing relationships between individuals. Participants were reimbursed £20 each to cover the costs of their travel and time. Updated ethical approval was acquired for the study (QMREC2013/48).

5.2.3 Setup

The experiment was held in the performance space at Queen Mary University of London, with overhead stage lighting, and black curtains surrounding the performance area. A large screen was positioned in the centre of the room, which showed the animation and provided instructions to the participants. The participants were positioned facing the screen, but angled slightly towards one another. The percussionist was provided with two electronic drum pads (snare and floor tom) and an electronic ride cymbal, which they played standing up. The pianist was provided with a 61 note keyboard and sustain pedal, which they also played standing up. Both instruments were connected to a computer via MIDI, and the audio was output through speakers positioned either side of the screen. Each participant's LuminUs was positioned on a stand in front of them, such that it was just below their line of sight to the screen. The devices were positioned such that each participant could only see their own device, and the brightness of the lights were adjusted so that no reflections were visible.

The ECG electrodes were attached to the participants' chests and the devices were strapped around their waists. The eye-tracking headsets were worn like normal glasses and attached to the computers using long USB extension leads; allowing the participants to move freely around the performance space. A video camera was discreetly set up just below the screen, so that it was facing the participants. Figure 5.2 shows the experimental setup as seen from this video camera.

Two iMac computers were positioned behind screens, so that the participants had their backs to them whilst they were playing. These computers were used to run the Pupil and LuminUs server software programs for each participant (see Section 4.4.2). They were also used by the participants to complete post-performance surveys. The experiment was controlled from a MacBook Pro, which was stationed outside of the performance area, behind the curtains. An additional Windows laptop was used to run the Shimmer software. Figure 5.3 shows a block diagram of the experimental setup. This diagram indicates how data were passed between the four computers and experimental hardware using both wired and wireless connections.

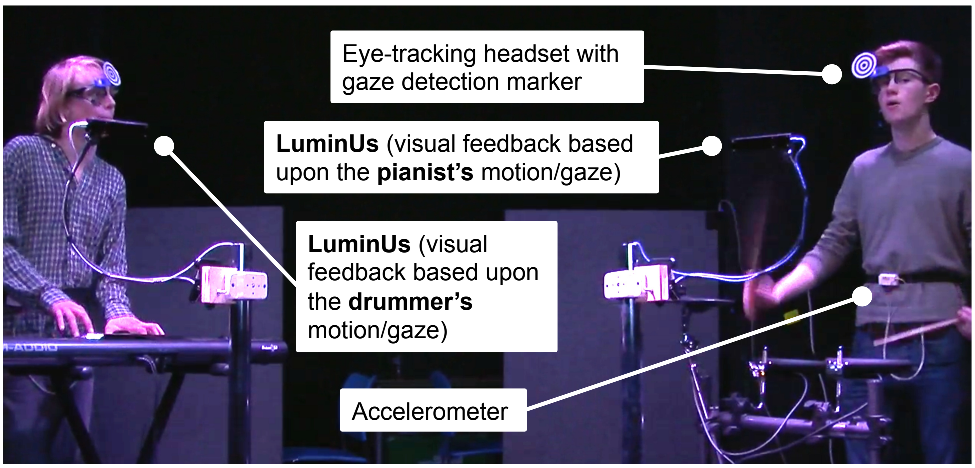


Figure 5.2: Annotated image of the experimental setup for Study 2.

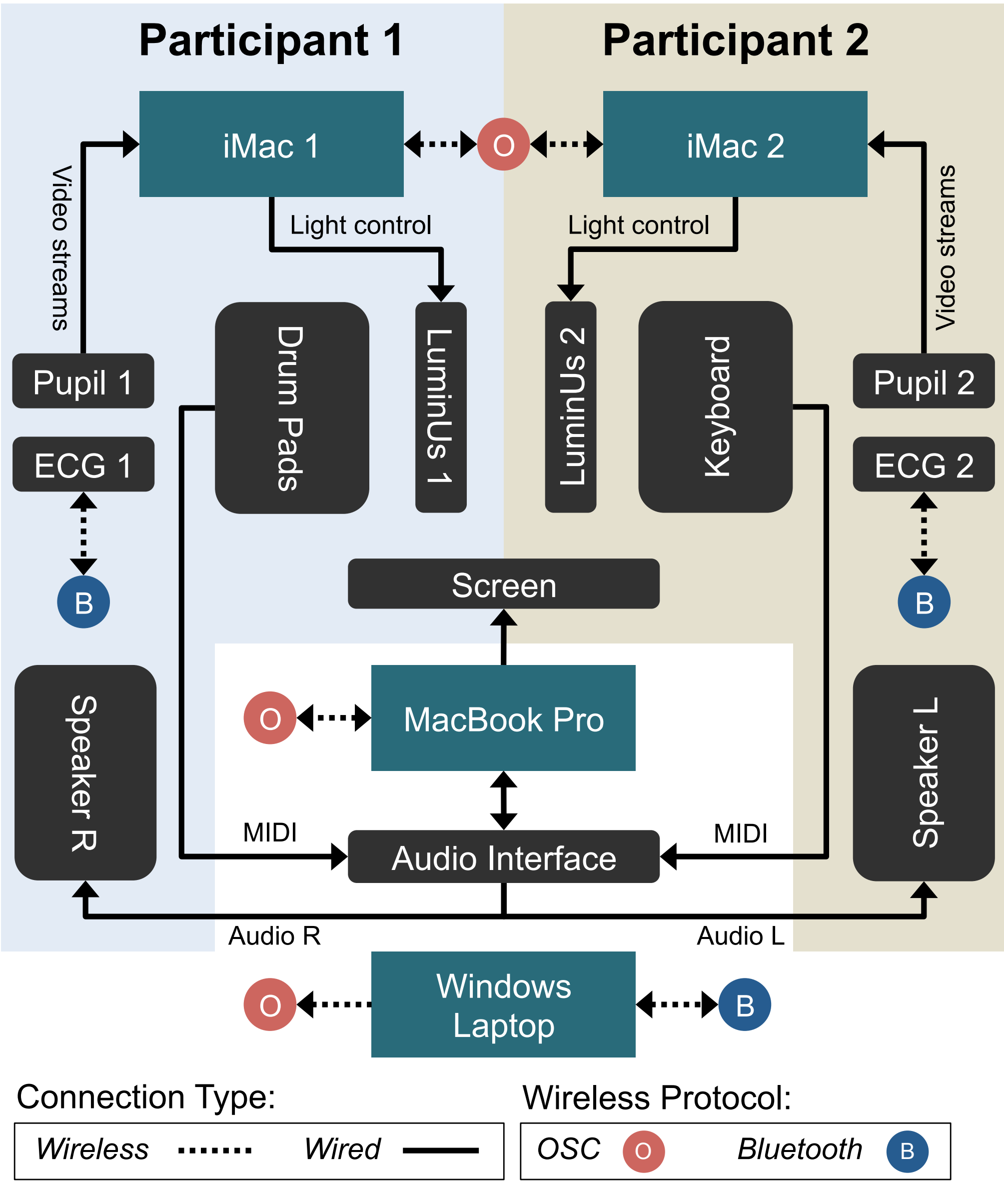


Figure 5.3: Block diagram of the experimental setup for Study 2.

5.2.4 Tasks

There were seven experimental conditions: G-G, G-X, X-G, M-M, M-X, X-M, and X-X. The two letters represent the feedback that the LuminUs provided to the percussionist and pianist respectively: either gaze feedback (G), motion feedback (M), or no feedback (X). The order of the experimental conditions was randomised at the start of each experiment. The experimental protocol was as follows:

1. **Arrival and introduction:** Participants were given verbal and written descriptions of the LuminUs and the experimental procedure. They were then asked to sign a consent form and were informed that they could leave at any point. Participants were told that in each condition they would either receive gaze feedback, motion feedback, or no feedback. They were then given a couple of minutes to play the instruments together. This provided an opportunity for them to make brief introductions and familiarise themselves with the instrumental set up.

2. **Sensor set up:** The participants were instructed on how to attach the ECG electrodes to their chests, which they did privately behind a screen. Following this, the eye-tracking headsets were set up and calibrated.

3. **Warm-up task:** The participants watched the animation twice without playing their instruments and twice whilst playing along. This allowed them to familiarise themselves with the animation and practice creating accompaniments. They were then given roughly one minute to discuss ideas. This was the last opportunity for them to interact verbally during the experiment.

4. **Improvisation task:** Participants were given two attempts to create an accompaniment to the animation. At the start of the task the screen instructed the participants to look at their LuminUs, which would then show either green-red (motion feedback), blue-purple (gaze feedback), or no lights (no feedback) to indicate which type of feedback they would be receiving. This was done so that neither participant would know what kind of feedback the other was receiving. For each attempt the screen provided a 10 second countdown to the start of the animation.

5. **Post-session survey:** The participants were instructed to go to the computers behind them and complete a short questionnaire. The questions were presented in a randomised order. The participants were not allowed to speak to each other during this time.

Repeat: Steps 4 and 5 were repeated a further six times. The feedback provided by the LuminUs changed for each improvisation task, based upon the randomly assigned conditions.

5.2.5 Measures

MIDI data were recorded from the keyboard and electronic drum kit, and the performances were also recorded on video. The ECG and accelerometer data from the Shimmer sensors were recorded at a sample rate of 512 Hz. Continuous eye-tracking data were recorded for each participant, consisting of the distance between the gaze point and the marker (if detected); and a binary number representing whether the participant was glancing at the marker (0 = no, 1 = yes). In addition to this, the animation playback time and a numerical representation of the light output from each LuminUs device were also recorded. All data were synchronised to a common timeline.

During the study subjective feedback was collected using a modified version of the post-performance questionnaire from the first study. The questions were as follows:

- On a scale of 0 to 10, how creative was each attempt?
- Which attempt do you think produced the best accompaniment to the animation? *[9-point slider ranging from 'Attempt 1' to 'Attempt 2', with 'both were equal' in the middle]*
- Regarding the entire last session (both attempts), how much do you agree/disagree with the following statements? *[7-point Likert scale labelled from 'strongly disagree' to 'strongly agree']*
  - The other musician was leading the performance.
  - I enjoyed this session.
  - The other musician and I performed well as a pair.
  - I was satisfied with my musical contribution.
  - I felt engaged with the other musician.
  - The session was boring.
  - The other musician ignored my contributions.
  - I liked the music we created.
  - I felt connected to the other musician.

5.2.6 Software

Custom software was written in order to control the experiment and record all the experimental data (see Fig. 5.4). The software was based upon the LuminUs control software (see Section 4.4.2), with the following additional capabilities:

Visualisations: Built in graphical displays enabled the accelerometer and ECG data to be visualised in real-time. This was done so that the researcher could verify that the sensors were working correctly throughout the experiment. A video display window was also included to enable the researcher to see what was currently being displayed on the participants' screen.

Control of the experiment: The software was programmed to run the experiment from the warm up task onwards (see Section 5.2.4). This involved assigning randomised experimental conditions; providing on-screen instructions and count-downs for the participants; playing the animation at the appropriate times; and sending the appropriate light control data to both of the LuminUs devices. The only input required from the researcher during this time was to click 'start' when the participants were ready to begin each improvisation task. A 'pause/resume' button was also included to allow the experimenter to pause the experiment in between tasks if there were any problems.

Data recording: The software received and synchronised the MIDI, ECG, accelerometer, eye-tracking, video playback, and LuminUs light control data. For each participant, two text files were created: one for the Shimmer data, and one for the Pupil data. A further text file was created containing the animation playback times and the corresponding LuminUs light control values (see Section 4.4.2). In each of these files, data were saved as comma-separated values; where the first three columns contained values representing the experiment run time (in ms), the experimental condition (0-6), and the improvisation task attempt (1 or 2). A log file was also created containing the experiment ID, the order of the randomised conditions, and any notes that were made by the researcher during the experiment. MIDI data were recorded to a single MIDI file.

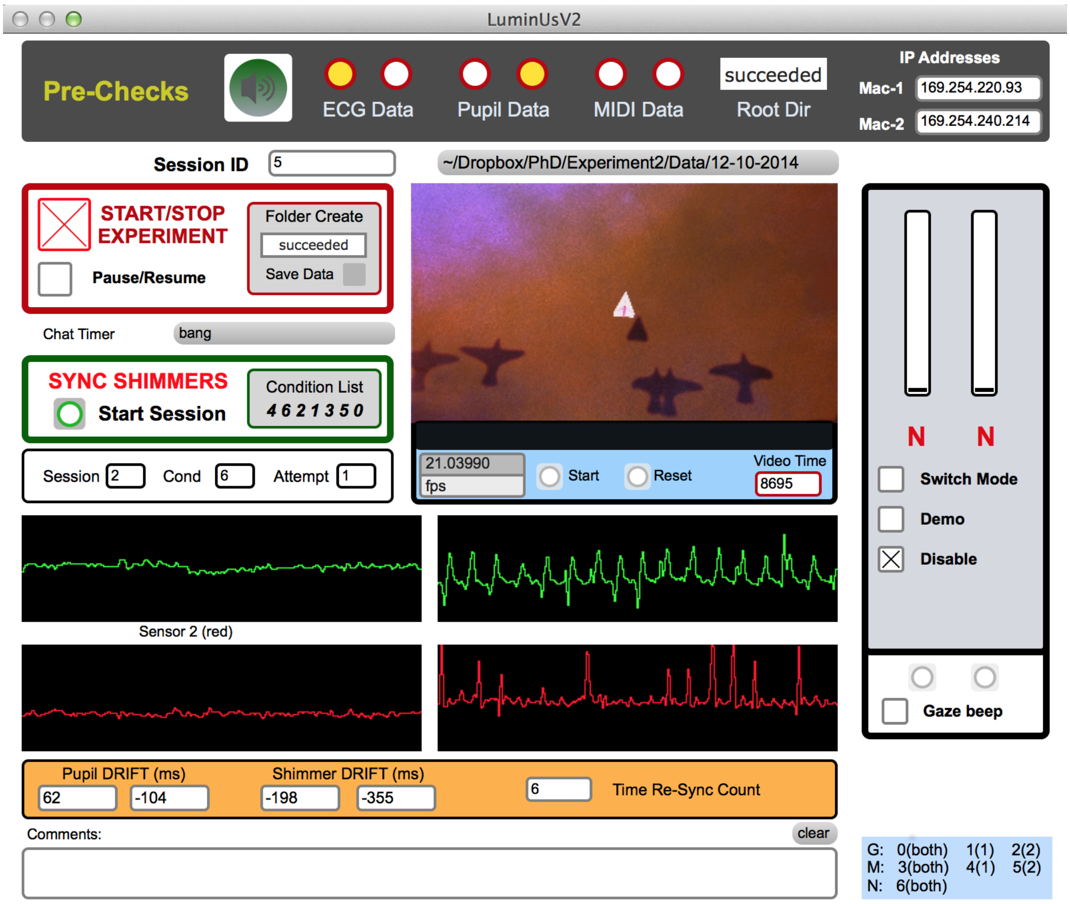


Figure 5.4: The control software graphical user interface (GUI) for Study 2.

5.3 Data Processing

Specific features were extracted from the data collected during the study. These features were chosen to test the experimental questions and hypotheses set out at the beginning of this chapter, and were informed by the research reviewed in Section 4.2. In this section the features and their methods of extraction are described.

In contrast to Study 1, which required data to be manually synchronised post-hoc, the data in the current study were all synchronised prior to being saved. All the data points were also labelled with a common time reference, and numbers identifying the experimental condition, and improvisation attempt. This meant that much of the extraction and processing could be automated. The resulting features were placed in feature tables, where the first four columns represented the experimental session ID (1-15); the participant (1 or 2); the experimental condition (0-6); and the improvisation attempt (1 or 2). Most of the data were processed in MATLAB using custom scripts.

5.3.1 Gaze Features

The start and end times of every detected glance (participant looking at their co-performer) were extracted from the Pupil data. These times were then used to calculate the duration of each glance, resulting in a feature table containing the individual glance timings (start and end) and durations for the entire experiment. This table was then used to compare the glance timings for pairs of participants within dyads, in order to detect when participants were looking at each other at the same time (mutual glance). A further table was created containing the start and end times, and durations of all the mutual glances. These basic tables were then used to create a statistical feature table, containing the following features for each participant and accompaniment attempt: the number of glances; time spent glancing; mean glance duration; number of mutual glances; time spent mutually glancing; and the mean duration of mutual glances.

5.3.2 ECG Features

Individual timings of detected heart beat events were extracted using the dedicated MATLAB scripts provided in ECGtools; similar to the ECG processing performed during Study 1 (see Section 3.3.2). An additional ECGtools script was used to remove any artifactual beats, based upon methods described by Clifford et al. (2002). A basic feature table was then created, where each row contained the timing of an individual heart beat; the corresponding time interval between this beat and the previous beat, known as the inter-beat interval (IBI); the heart rate (HR); and a number representing whether that point was a peak/maximum (1), trough/minimum (-1), or neither (0). This basic table was then used to create a statistical feature table, containing the following features for each participant and accompaniment attempt: mean IBI time; standard deviation in HR; and the number of HR extrema.

Heart rate variability (HRV) features were also extracted using the open-source heart rate variability analysis software (HRVAS) (Ramshur, 2010), which also runs in MATLAB. The following commonly used HRV features (described in detail in Section 4.2.3) were extracted from the ECG data and included in the aforementioned statistical feature table: SDNN; LF (low frequency) band power; HF (high frequency) band power; and LF/HF ratio. An additional time-frequency domain feature was extracted, consisting of a continuous representation of the LF/HF ratio over the course of a given recording period. This was performed in HRVAS using the wavelet transform method with a window size of 30 seconds, and an overlap of 15 seconds, as described by Ramshur (2010). The recording periods consisted of individual accompaniment attempts.

5.3.3 Accelerometer Features

The raw accelerometer data consisted of the summed absolute X, Y, and Z axial accelerometer components. These were initially down-sampled, due to the fact that they had been sampled at the same rate as the ECG data (512 Hz). To remove the effects of gravitational acceleration, the differences were then calculated between consecutive down-sampled values; giving the rate of change of acceleration (jerk). A Savitzky-Golay FIR filter was then used to smooth these data. The resulting values were placed into a basic feature table, which was then used to create a statistical feature table, containing the following features for each participant and accompaniment attempt: mean body motion (jerk); and standard deviation in body motion.

5.3.4 Decision Making Features

Three features were extracted to provide quantitative measures of the participants' musical decision making. The first two of these derived from the MIDI data, whilst the third was based upon the animation used in the study. These features were used to analyse the effects of the LuminUs upon decision making, as per question **LQ4** in Section 5.1.1; and to analyse relationships between cardiac activity and musical decision making, as per the hypotheses in Section 5.1.2.

MIDI-derived Features

In the first study rhythmic change points were manually extracted from the MIDI data, which was a time consuming process. The present study consisted of a total of 840 minutes of MIDI data (15 dyads $\times$ 2 participants $\times$ 7 conditions $\times$ 2 attempts $\times$ 2 minutes per attempt). Therefore, it was not feasible to manually label musical changes in the data. Furthermore, manual labelling is prone to subjective bias. Consequently, automatic methods of extracting musical change features from MIDI data were sought. The two selected methods are described below. It should be noted that the MIDI data were found to contain a timing drift, due to software issues. This meant that over the course of each study session the MIDI data became out-of-sync with the rest of the data. This was accounted for by automatically approximating the drift and applying uniform scaling to the MIDI time values. However, due to this being an approximation, the resulting MIDI note timings had an estimated error of ± 5 seconds.

Information Content: Pearce (2005) developed and evaluated a computational model of melody perception: the Information Dynamics of Music (IDyOM) model[2]. Through exposure to melodies, the IDyOM model learns about sequential dependencies between notes. Unsupervised learning and prediction are performed using a Markov, or n-gram model (Manning and Schütze, 1999), which is trained on existing melodic sequences. The result is that the model can compute the conditional probability of a note, given the prior notes in that melody. The negative log probability is then used to give the information content (IC), which is a representation of the unexpectedness of the note (Wiggins et al., 2009). The IC rises when a melodic change occurs; therefore, this feature can be used to predict structural changes in a piece of music (Potter et al., 2007). Consequently, the IDyOM model was used to automatically quantify musical change in the MIDI data from the present study. In order to do this, the MIDI data from each experimental session were segmented to create MIDI files (compositions) for each accompaniment and each participant. The batches of 14 compositions (7 conditions × 2 attempts) for each participant were then imported into IDyOM as datasets.

IDyOM allows the user to specify whether the training is performed using a short-term model (STM), which only uses the current composition; a long-term model (LTM), which uses a set of compositions; or a model combining both the STM and LTM (BOTH). The latter was used for this study, meaning that, for each composition in a dataset, training was based upon that particular composition, as well as all the other compositions performed by that participant.

Additional parameters that must be selected in order to run IDyOM are the target and source *viewpoints*. The target viewpoints are the features of the notes (e.g. pitch, duration) that the model attempts to predict; whilst the source viewpoints are the features that are used for performing the prediction. Different viewpoints were used depending upon whether the MIDI data were from a drummer or a pianist. For drum MIDI data the target viewpoint was the note onset timing, and the source viewpoint was the inter-onset-interval ratio (the current time interval between notes divided by the previous time interval between notes). For piano data the target viewpoint was the pitch, and the source viewpoint was the pitch interval.

Following the processing of all 15 datasets, the resulting output data from IDyOM were imported into a single feature table, containing the note timing and IC value for every MIDI note. Once again, these values were used to create a statistical feature table, containing the following features for each participant and accompaniment attempt: mean IC; standard deviation in IC.

[2] `https://code.soundsoftware.ac.uk/projects/idyom-project`

Change Points: IDyOM has predominantly been designed and evaluated for use with monophonic melodies (Pearce, 2005), so its performance with polyphonic and rhythmic data is not well established. Consequently, an additional method was developed for extracting decision making features from the MIDI data. The method attempts to predict where musical changes are occurring in a given composition by using three components of the MIDI notes: timing (nT), pitch (nP), and velocity (nV). Each of these components is used to compute a time series feature. This is achieved by stepping through the MIDI timeline in fixed time increments and calculating the three features at each step. In each case, the feature calculation is performed by analysing the extent to which the set of notes leading up to that point differ from those following it. If the difference is high then it suggests that a musical change may have occurred.

The subscripts $_F$ and $_P$ are used to refer to future notes and past notes respectively. For example, the term nT_F refers to all of the timing values for the future notes; whilst nP_P refers to the pitch values of all past notes. In each case, the sets of future and past note components for a time point t_s are given by:

$$\begin{aligned} nX_F &= nX(t) \quad for \quad \{t \mid t_s \leq t < (t_s + t_w)\} \\ nX_P &= nX(t) \quad for \quad \{t \mid (t_s - t_w) \leq t < t_s\} \end{aligned} \tag{5.1}$$

Where t_w is a given time window, and $nX(t)$ is the set of all components at the times specified by t. The timing feature is calculated as:

$$\begin{aligned} f_{time}(t_s) = \frac{1}{n_p} \sum_{p=21}^{108} |std\Big[\text{diff}\big(nT_F(p)\big)\Big] - std\Big[\text{diff}\big(nT_P(p)\big)\Big]| \\ for \quad \{p \mid p \in nP_F \wedge p \in nP_P\} \end{aligned} \tag{5.2}$$

Where $nT_F(p)$ is an array containing all the timings of the future MIDI notes with MIDI pitch value p. The standard deviation (std) of the set of differences (diff) between adjacent note timings is calculated separately for past and future components; then the absolute difference between these standard deviations is found. This difference is then calculated and summed over all pitch values, where the pitch value must be contained within both of the sets: nP_F and nP_P. Finally, the value is divided by the total number of unique pitch values (n_p). In summary, this feature looks at a representation of the spacing between past and future notes and uses it to predict whether timing changes have occurred.

The pitch feature is calculated as:

$$f_{pitch}(t_s) = \frac{size\big[\text{setdiff}(nP_F, nP_P)\big]w_F}{size(nP_F)} + \frac{size\big[\text{setdiff}(nP_P, nP_F)\big]w_P}{size(nP_P)} \tag{5.3}$$

Where setdiff(nP_F, nP_P) gives the set of pitch values that occur in future notes (nP_F) but do not occur in past notes (nP_P); and vice-versa for setdiff(nP_P, nP_F). The size of this set is given by the *size* operator, therefore $size(nP_F)$ is simply the number of notes in the future set. The weight, w_F, is the number of notes in nP_F with pitch values that do not occur in nP_F; and vice-versa for w_P. In summary, this feature looks at the number of unique note pitches in the past and future sets, and scales these values according to proportion of new notes out of all the notes played. A large difference between the notes played in the past and those played in the future will be indicative of a musical change.

Finally, the velocity feature is calculated simply as:

$$f_{velocity}(t_s) = |mean(nV_F) - mean(nV_P)| \tag{5.4}$$

Where $mean(nV_F)$ is the mean note velocity for all future notes in the set nV_F. This feature simply consists of the absolute difference between the mean velocity of the future set and the mean velocity of the past set. If there is a big change in velocity, then this may be indicative of musical change.

In order to extract discrete change points (CPs), a weighted sum of the three time series features is created. Peaks in this combined series are then extracted to provide the CPs. Peaks that are below a given threshold are discarded; and for any peaks that are less than four seconds apart the lower peak is discarded. The weights for each component feature were determined by selecting a random sub-set of accompaniments and manually evaluating the performance of the CP detection for different weights. Different weights were selected for drum data and piano data. This is due to the fact that certain features are more powerful in each case. For example, the pitch feature is not particularly useful for detecting drum CPs.

Figure 5.5 shows an example of the CP detection, including the three time series features, the detected CPs, and the original MIDI data. Once all the data were processed, the CP time values for every accompaniment were extracted into a single table. These values were used to create a statistical feature table, containing the following features for each participant and accompaniment attempt: number of CPs; mean time between CPs; standard deviation in the time between CPs.

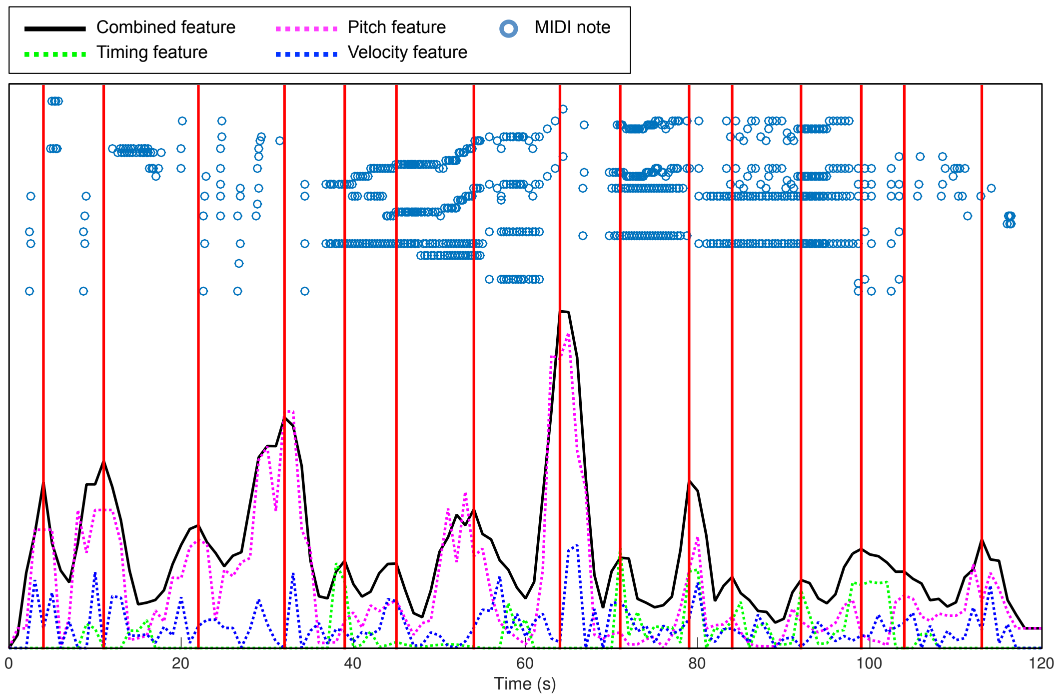


Figure 5.5: Example plot of change point features and MIDI notes. Detected change points are indicated by the vertical lines.

Animation Cues

As previously discussed, one of the purposes of the animation was to introduce visual cues that might stimulate the musicians to change their playing at similar points (see Section 5.2.1). Consequently, the animation was analysed in order to identify distinct animation cue points. This was primarily achieved through the manual identification of points in the animation where changes occurred. This was carried out by the researcher and a professional musician, independently. There was a high level of agreement, with a total of 7 cue points identified, six of which were mutually agreed upon. Descriptions of the resulting cue points are given in Table 5.3. In addition to this, a quantitative indicator of image-based changes in the animation was also used. This involved plotting the frame difference between consecutive frames in the animation. As can be seen in Fig. 5.6, the resulting phases and peaks in the frame difference plot showed close matches to the subjectively identified cue points. Therefore, no further changes were made to the set of 7 subjectively identified cue points.

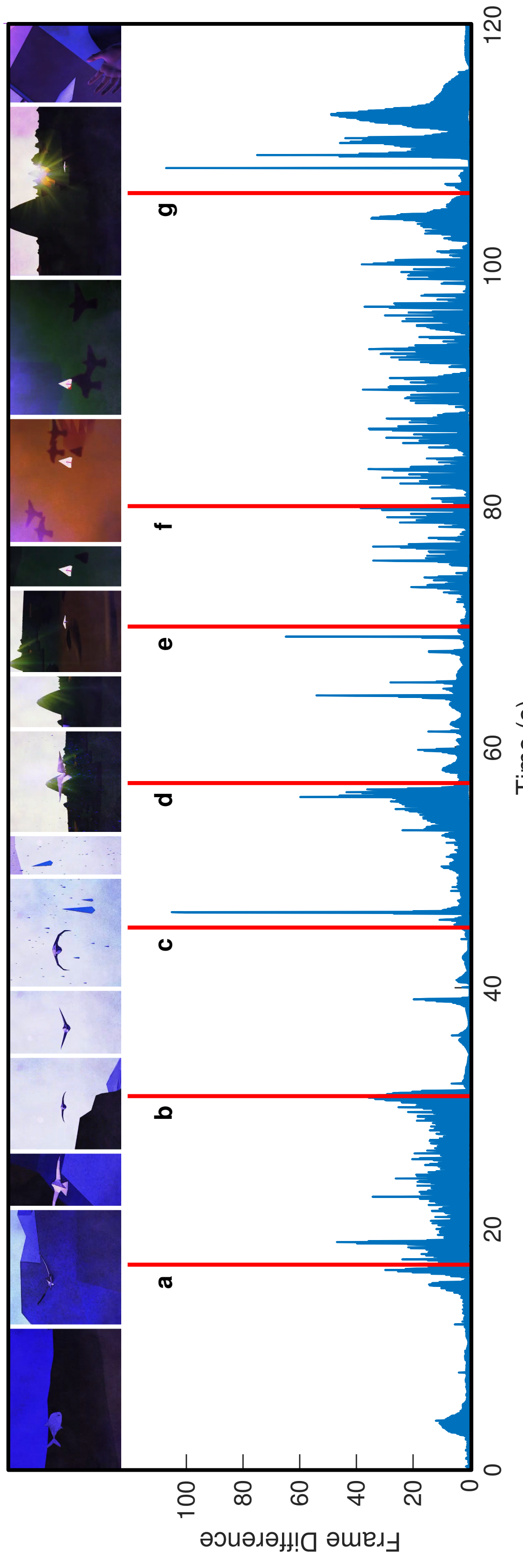


Figure 5.6: Plot of the animation frame difference with cue points (vertical lines) and still images.

Table 5.3: Animation cue point descriptions.

Cue	Time (s)	ID	Description
a	17	R M	The fish swimming in the stream becomes a bird.
b	31	R M	The bird flies up above the stream and into the sky.
c	45	R M	Rain starts falling and the bird turns into paper.
d	57	R M	The paper becomes a paper plane and starts to fly.
e	70	R M	The camera switches view from behind to overhead.
f	80	R	Bird shadows appear on the ground below.
g	106	R M	The camera view switches back to behind, and the end sequence begins.

Note: R = Identified by researcher, M = Identified by musician.

5.3.5 LuminUs Feedback Features

LuminUs feedback features were extracted as quantitative measures of the amount of feedback that the participants received during each accompaniment. Two features were calculated: L-time, and L-persec. L-time is simply the amount of time that the LuminUs was lit up during the accompaniment. L-persec is a measure of how many LEDs were lit per second, on average. It is calculated by summing the number of lit LEDs at each time point during an accompaniment, and then dividing this by the total number of time points.

5.4 Analyses and Results: Effects of the LuminUs Feedback

The exploratory questions set out in Section 5.1.1 concern the analysis of the effects that the LuminUs had on the behaviours and experiences of the musicians. In this section various statistical analyses are used to investigate the effects of the LuminUs on i) **gaze behaviour**; ii) **body motion**; iii) **musical decision making**; and iv) **self-report** measures of the collaborative performances.

5.4.1 Gaze Behaviour

As a starting point for the analysis of effects of the LuminUs upon gaze, three features were analysed: i) number of glances; ii) time spent glancing; and iii) glance duration. These features provide information on the mean glancing behaviour of the participants during each accompaniment session. A glance is defined as the specific act of looking towards the head of one's co-performer. The bar graphs in Fig. 5.7 show the mean and standard error for the three glance features, averaged over participants within each

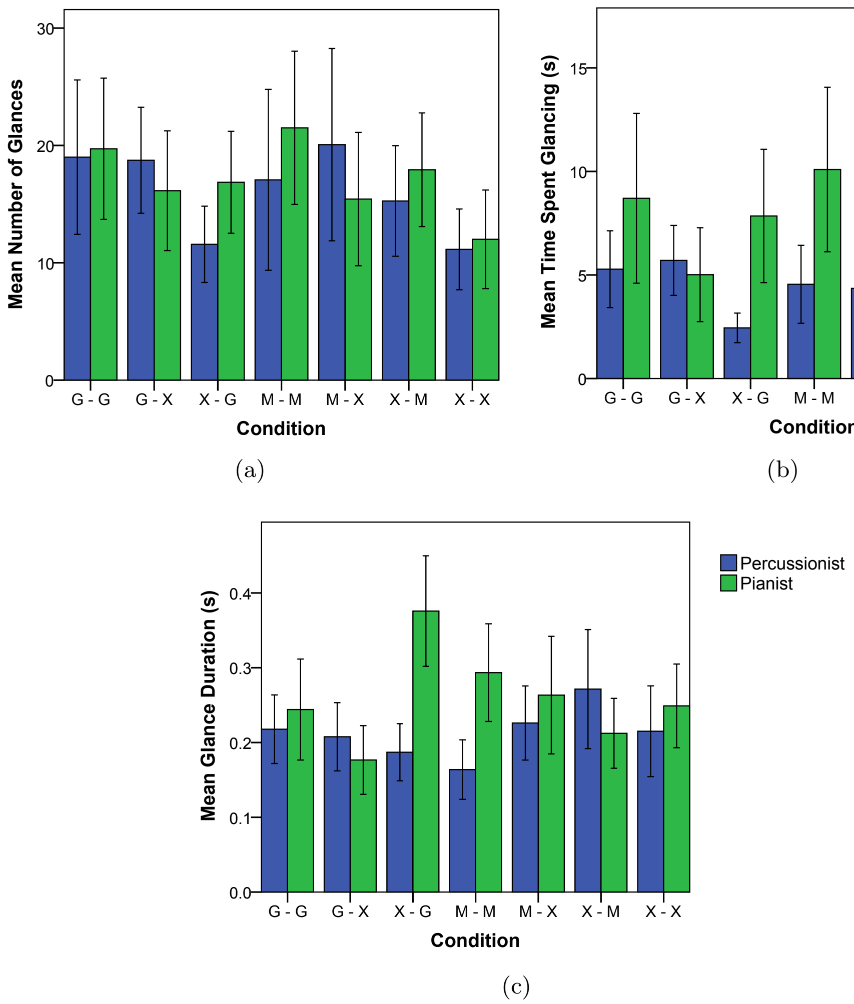


Figure 5.7: Bar graphs showing (a) the mean number of glances; (b) the mean time spent glancing; and (c) the mean glance duration within each of the seven conditions. G = gaze feedback, M = motion feedback, X = no feedback. Error bars: ± 1 SE.

condition. Figure 5.7(a) shows that the mean number of glances for both the pianist and percussionist is generally highest in the conditions where the LuminUs feedback was enabled. This trend appears to exist regardless of whether the feedback being provided is motion feedback or gaze feedback. Figure 5.7(b) shows that the time spent glancing is also generally higher in the conditions where LuminUs feedback is enabled. However, for the motion feedback conditions this trend is not as uniform. This can be seen by comparing the results for the two conditions where only one musician is receiving motion feedback (M-X and X-M), where, in each case, the musician receiving feedback glances for less time relative to the condition where only their partner receives motion feedback. Figure 5.7(c) shows the average duration of glances in each condition.

In this case, there does not appear to be any consistent trend. Specifically, the glance durations in conditions where LuminUs feedback is enabled are not notably higher or lower than in the conditions where feedback is disabled. An exception to this is the result for the mean pianist glance duration in condition X-G. This value is substantially higher than in most of the other conditions. A potential explanation for this lies in fact that each musician was only aware of their own feedback condition. This means that one or more pianists could have been under the impression that the percussionist was also receiving gaze feedback in this condition. Consequently, these pianists might have expected their glances to trigger a response from the percussionist, prompted by the LuminUs feedback. If no response was forthcoming then this might have led the pianists to hold their glances for longer.

For each of the three features plotted in Fig. 5.7, Friedman tests were conducted to assess whether differences existed between the seven conditions. There was a statistically significant difference in the number of glances, depending on which type of LuminUs feedback was being provided ($\chi^2(6) = 19.264$, $p = .004$). No significant difference was found for the mean time spent glancing ($\chi^2(6) = 7.971$, $p = .240$), nor for the mean glance duration ($\chi^2(6) = 6.134$, $p = .408$). Consequently, the Wilcoxon signed rank test was conducted to analyse the differences between the mean number of glances obtained in the six conditions where the LuminUs was enabled, relative to the equivalent means for the inactive LuminUs condition (X-X). Due to multiple comparisons, a Bonferroni correction[3] was applied, resulting in a significance level set at $p < .008$ (). The results in Table 5.4 indicate that, relative to the no-feedback condition

[3] The Bonferroni correction is used to avoid Type I errors when making multiple comparisons. It is calculated by dividing the significance level (in this case 0.05) by the number of tests that are being performed.

Table 5.4: Statistical results for the differences between the number of glances in conditions 1-6 (LuminUs in use), relative to condition 7 (LuminUs not in use). Only data for participants receiving LuminUs feedback in each condition were used.

Condition	Z	r	n	p
1) G-G	**3.81***	0.70	27	.000
2) G-X	**2.75***	0.71	14	.006
3) X-G	2.12	0.55	14	.034
4) M-M	**2.83***	0.52	28	.005
5) M-X	1.92	0.50	14	.054
6) X-M	2.17	0.56	14	.030

Note: Z = Wilcoxon signed rank z-statistic, r = effect size, n = sample size, p = two-tailed p-value. $^*p < .008$ (Bonferroni correction applied).

(X-X), there were significantly more glances in the condition where both participants had gaze feedback (G-G) ($r = 0.7$, $p = .000$), and where only the percussionist had gaze feedback (G-X) ($r = 0.7$, $p = .006$). Looking at the motion feedback conditions, there were significantly more glances in the condition where both participants are receiving feedback (M-M) ($r = 0.52$, $p = .005$).

The previous results indicate that the LuminUs feedback had significant effects upon the number of glances made by participants for whom it was enabled. For the condition where only percussionists were receiving gaze feedback (G-X) there was an interest in whether the observed effects also led to knock-on (secondary) effects on the number of glances made by the pianists. To investigate this, a Wilcoxon test was conducted for differences between the mean number of glances in condition 2 (G-X), relative to the no-feedback condition (X-X), but only using the data for pianists. The results showed no significant differences ($r = 0.37$, $p = .154$).

The analyses above look exclusively at the effects of the LuminUs upon the glancing behaviours of individuals. They do not investigate the effects of the LuminUs on the gaze interactions *between* musicians. Specifically, there was an interest in how the LuminUs may have influenced the fluency of visual exchanges between individuals in each dyad. In order to investigate this, an analysis of the effects of the feedback conditions upon *glance reciprocation* was performed. A reciprocated glance is defined as one that occurs within a given time delay of the start of a prior glance towards that individual. For example, a pianist's glance towards their fellow percussionist would be considered as a reciprocated glance if the percussionist glanced back within five seconds. The glance timing data were used to count the proportion of each participant's glances that were reciprocated within each condition. No consensus was found in the literature regarding the length of the delay for which a glance can be considered to be reciprocated. Therefore, the glance reciprocation features were extracted using a range of time delays from 5 seconds to 1 second. Friedman tests revealed significant effects of the feedback condition upon glance reciprocation for a time delay of 5 seconds ($\chi^2(4) = 10.407$, $p = .034$). No significant differences were found for delays of 4 seconds ($\chi^2(4) = 8.757$, $p = .067$), 3 seconds ($\chi^2(4) = 5.061$, $p = .281$), 2 seconds ($\chi^2(4) = 6.377$, $p = .173$), and 1 second ($\chi^2(4) = 2.196$, $p = .700$). Consequently, the Wilcoxon signed rank test was used to compare the proportion of glances reciprocated with 5 seconds in each of the six conditions where the LuminUs was enabled, relative to the inactive LuminUs condition (X-X). Again, a Bonferroni correction was applied, resulting in a significance level set at $p < .008$. The results in Table 5.5 indicate significant differences between condition 1 (G-G) and condition 7 (X-X) ($r = 0.61$, $p = .008$). This suggests that the

Table 5.5: Statistical results for the differences between the proportion of glances reciprocated within 5 seconds in conditions 1-6 (LuminUs in use), relative to condition 7 (LuminUs not in use). Only data for participants receiving LuminUs feedback in each condition were used.

Condition	Z	r	n	p
1) G-G	**2.65***	0.61	19	.008
2) G-X	2.24	0.71	10	.025
3) X-G	0.85	0.27	10	.398
4) M-M	1.81	0.40	20	.070
5) M-X	1.13	0.34	11	.260
6) X-M	−0.70	−0.22	10	.484

Note: Z = Wilcoxon signed rank z-statistic, r = effect size, n = sample size, p = two-tailed p-value. $^{*}p < .008$ (Bonferroni correction applied).

receipt of gaze feedback caused participants to reciprocate significantly more of their co-performer's glances. Interestingly, this effect did not reach significance for conditions 2 and 3, where gaze feedback was only provided to one of the two musicians. However, the effect size for the condition where only percussionists received gaze feedback (G-X) ($r = 0.71$, $p = .025$) is notably greater than for the condition where only pianists received gaze feedback (X-G) ($r = 0.27$, $p = .398$). This suggests that the effects of the LuminUs may be dependent upon the instrument being played.

The previous statistical analyses looked at the effects of the LuminUs feedback upon gaze behaviour by analysing differences between experimental conditions. These analyses treated the level of feedback being provided by the LuminUs as binary - either it is enabled, or it is disabled. This approach is valid, especially when considering that the participants' knowledge of the experimental condition alone may have influenced their behaviours. However, it was also possible to undertake more detailed analyses using quantitative measures of the feedback provided by the LuminUs in each condition. These analyses were facilitated by the fact that the exact lighting output of the LuminUs was recorded during the experiments. In order to perform these analyses, a linear mixed model (LMM) approach was adopted, similar to that used in Section 3.4. LMMs were required in order to perform valid statistical tests on relationships between quantitative measures of LuminUs feedback and glance behaviour, using multiple non-independent samples from individual participants. Two features were used, which quantitatively represent the amount of feedback provided by the LuminUs in slightly different ways (see Section 5.3.5). The first feature is the L-time, which is simply the amount of time that the LuminUs was illuminated during each condition. The second feature is L-persec, which represents the average number of LEDs that were illuminated during

Table 5.6: Linear mixed effects modelling (LMM) estimates of fixed effects of two measures of the light feedback provided by the LuminUs (L-time and L-persec) upon the number of glances. Only data for participants receiving LuminUs feedback in each condition were used.

Parameter	Estimates of Fixed Effects				
	Estimate	Std. Error	df	t	p
Gaze feedback:					
L-time	0.000613	0.000222	22.95	**2.766***	.011
L-persec	6.661241	2.849823	33.00	**2.337***	.026
Motion feedback:					
L-time	−0.001388	0.000654	15.52	**−2.124***	.050
L-persec	−1.072314	1.746276	114.54	−0.614	.540

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. Significance estimates: $^{*}p < .05$. All estimates were calculated using *unstructured* covariance type for random effects. *Scaled identity* and *diagonal* repeated covariance types were used for gaze feedback and motion feedback estimates respectively.

Table 5.7: Linear mixed effects modelling (LMM) estimates of cross-participant effects of the number of glances.

Dependent Variable	Parameter	Estimates of Fixed Effects				
		Estimate	Std. Error	df	t	p
P1 Glance count	P2 Glance count	0.198092	0.147973	15.47	1.339	.200
P2 Glance count	P1 Glance count	0.254155	0.105949	6.26	2.339	.052

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. All estimates were calculated using *variance components* covariance type for random effects and *diagonal* covariance type for repeated measures.

each second. The analyses in Table 5.4 indicated that both gaze and motion feedback conditions had a positive influence upon the number of glances that the musicians made. Consequently, there was an interest in whether corresponding relationships existed between the amount of feedback provided by the LuminUs and the number of glances made. A LMM approach was used to model the effects of the two LuminUs feedback features (L-time and L-persec) upon the number of glances. This analysis was performed for each of the two types of feedback - gaze and motion. The results are provided in Table 5.6. For gaze feedback, both of the LuminUs feedback features have significant positive effects upon the number of glances (L-time: $t = 2.766$, $p = .011$; L-persec: $t = 2,337$, $p = .026$). For motion feedback there is a significant negative effect of L-time on the number of glances ($t = -2.124$, $p = .050$). This latter result is interesting, since it contrasts with the earlier results (Table 5.4), which indicated that, much like the glance feedback condition, the motion feedback condition led to increased glancing. In light of the results in Table 5.6, it is clear that there are distinct differences between the ways in which the gaze and motion feedback affected the number of glances. This implies that the participants specifically responded to the gaze feedback by glancing more towards their co-performer. However, given that the LuminUs's gaze feedback is a direct representation of the co-performer's glances, it is necessary to consider the possibility that these results simply reflect the influence that one participant's glances had upon the other participant. This influence could be achieved through peripheral vision alone. To test whether there was an underlying correlation between the glances of each participant in a dyad, a LMM analysis was performed on cross-participant effects of the number of glances. The results in Table 5.7 show that, although there are positive effects in each case, the effects are not significant. These results address question **LQ3** and indicate that the gaze feedback provided by the LuminUs was, in itself, a positive influence on the number of glances exchanged during the musical interactions.

5.4.2 Body Motion

Much like the analysis of gaze behaviour in the previous section, the analysis of motion behaviour was commenced by graphing the mean body motion of the percussionists and pianists in each of the experimental conditions. The results are shown in Fig. 5.8. It is immediately obvious that, on average, the percussionist moved more than the pianist, as might be expected. When comparing the bars for conditions where the LuminUs was enabled to those where it was not, there do not appear to be any consistent or noteworthy trends. Looking specifically at the motion feedback conditions, it appears that there is generally less motion than the other conditions, especially for the percussionist.

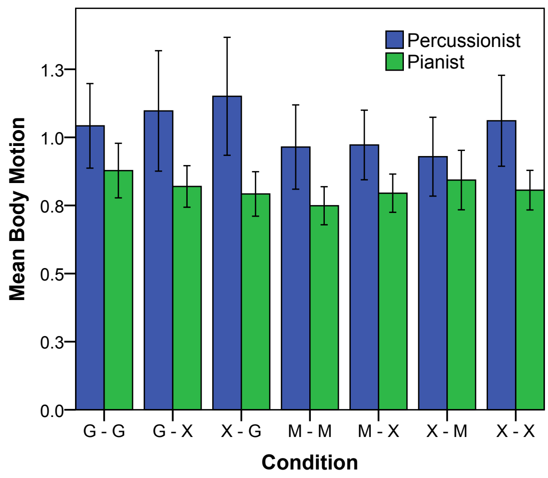


Figure 5.8: Bar graph showing the mean pianist and percussionist body motion within each of the seven conditions. G = gaze feedback, M = motion feedback, X = no feedback. Error bars: ± 1 SE.

A repeated measures ANOVA with a Greenhouse-Geisser correction was conducted, indicating no statistically significant differences in body motion across the seven conditions ($F(3.526, 102.226) = 1.277$, $p = .285$). Again, a LMM approach was used to analyse the effects of the amount of LuminUs feedback upon the amount of body motion. The results in Table 5.8 indicate that there are no significant effects of gaze feedback on the mean body motion of the musicians. For the motion feedback there is a significant effect of the L-persec feature upon mean body motion ($t = 3.3594$, $p = .002$).

Similar to the previous section, it is necessary to consider the fact that the LuminUs feedback is directly related to the motion of the other participant. Therefore, the significant effect observed in Table 5.8 could simply be due to underlying correlations between the mean body motion of each participant in each dyad. To test the likelihood of this, a LMM analysis was performed on cross-participant effects of body motion. Table 5.9 provides the results of this analysis and shows that there are significant and strong correlations between the mean body motion of the two participants in a dyad. These results might be expected, since the body motion of the performers is likely to be closely related to the dynamics and tempo of the music they are playing together. The results of this analysis mean that it is unlikely that the motion feedback of the LuminUs, in itself, had significant effects upon the body motion of the performers.

Table 5.8: Linear mixed effects modelling (LMM) estimates of fixed effects of two measures of the light feedback provided by the LuminUs (L-time and L-persec) upon the body motion of the musicians. Only data for participants receiving LuminUs feedback in each condition were used.

Parameter	**Estimates of Fixed Effects**				
	Estimate	Std. Error	df	t	p
Gaze feedback:					
L-time	−0.000002	0.000005	78.26	−0.34	.736
L-persec	−0.000831	0.111256	45.15	−0.007	.994
Motion feedback:					
L-time	−0.000013	0.000023	84.13	−0.549	.584
L-persec	0.252083	0.075052	45.799	**3.359****	.002

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. Significance estimates: $^{**}p < .01$. All estimates were calculated using *unstructured* covariance type for random effects. *Scaled identity* and *diagonal* repeated covariance types were used for gaze feedback and motion feedback estimates respectively.

Table 5.9: Linear mixed effects modelling (LMM) estimates of cross-participant effects of the mean body motion (b-motion).

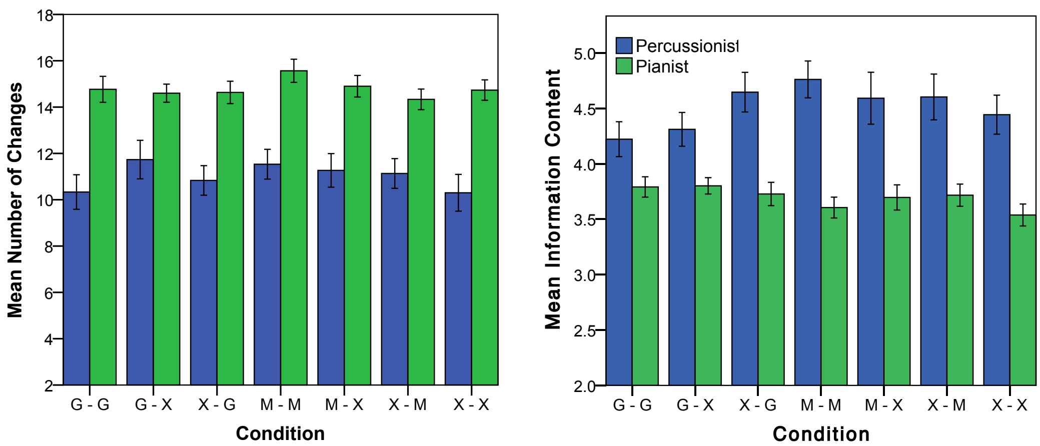
Dependent Variable	**Parameter**	**Estimates of Fixed Effects**				
		Estimate	Std. Error	df	t	p
P1 Mean b-motion	P2 Mean b-motion	0.625911	0.220671	16.76	**2.836***	.012
P2 Mean b-motion	P1 Mean b-motion	0.289424	0.078139	11.32	**3.704****	.003

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. Significance estimates: $^{*}p < .05$, $^{**}p < .01$. All estimates were calculated using *variance components* covariance type for random effects and *diagonal* covariance type for repeated measures.

5.4.3 Musical Decision Making

Exploratory question **LQ4** concerns the influence of the LuminUs feedback upon the musical decisions made by the musicians. As discussed in Section 5.3.4, two features were extracted from the MIDI data to provide quantitative measures of the musical decisions made by the musicians. The first feature consists of musical change points (CPs), which indicate points in each accompaniment where changes appeared to occur. These result in estimated values for the mean number of musical changes made by each participant in each accompaniment. The second feature is the information content (IC) of the music. This assigns an IC value to each note, based upon statistical models of musical structure. Once again, single features can then be extracted, representing the average IC for each participant and each accompaniment. Figure 5.9 shows the mean number of changes (5.9(a)) and the mean IC (5.9(b)) averaged over percussionists and pianists within each experimental condition. Clear differences can be seen between the mean values for the percussionists and the pianists; however, this is to be expected due to the different methods used to extract the features for each type of instrument. Looking across conditions, there do not appear to be any outstanding trends for the mean number of changes. For the IC, it appears that the percussionists have a lower mean IC in the conditions where they receive gaze feedback; whereas for the pianists, the mean is higher in these conditions relative to the no-feedback condition (X-X).

As in the previous analyses, Friedman tests were conducted to test for differences in the two musical decision making features across the seven feedback conditions. Unlike



(a) (b)

Figure 5.9: Bar graphs showing (a) the mean number of musical changes; and (b) the mean information content within each of the seven conditions. G = gaze feedback, M = motion feedback, X = no feedback. Error bars: ± 1 SE.

previous analyses, the statistical tests were conducted for pianists and percussionists independently, due to the difference in the methods used to extract the two features of interest - CPs, and IC. For the CP feature, no significant differences were found for percussionists ($\chi^2(4) = 5.732$, $p = .220$), nor for pianists ($\chi^2(4) = 4.954$, $p = .292$). For the IC feature, significant differences were found for percussionists ($\chi^2(4) = 16.373$, $p = .003$), whilst no significant differences were found for pianists ($\chi^2(4) = 5.493$, $p = .240$). Consequently, the Wilcoxon signed rank test was used to analyse differences between the percussionists' mean IC obtained in the four conditions where the LuminUs was providing them with feedback, relative to the no-feedback condition. The results are provided in Table 5.10. A Bonferroni correction has been applied, resulting in a significance level set at $p < .013$. The only significant difference is for condition 3 (M-M) ($r = 0.67$, $p = .009$). However, no corresponding result is obtained for the second condition where percussionists had motion feedback (X-M). This suggests that there was not any general influence of motion feedback upon the mean IC.

LMM analyses were performed in order to further analyse the potential effects of the LuminUs feedback upon the two measures of musical decision making; using L-persec as a quantitative measure of the amount of LuminUs feedback provided to each musician in each condition. The results are presented in Table 5.11. For the gaze feedback there are no noteworthy results. For the motion feedback there appears to be a significant negative effect of the amount of LuminUs feedback upon the mean IC for percussionists ($t = -3.289$, $p = .002$).

Given that the analyses above showed effects of both motion and gaze feedback upon IC, there was an interest in whether any underlying relationships existed between IC and the glances and body motions of the performers. To investigate this, LMM analyses were performed on the effects of both the mean body motion of the performers and the

Table 5.10: Statistical results for the differences between the mean information content for percussionists within the conditions where they received LuminUs feedback, relative to the no-feedback condition.

Condition	Z	r	n	p
1) G-G	−1.42	−0.37	15	.156
2) G-X	−1.19	−0.31	15	.233
4) M-M	**2.61***	0.67	15	.009
5) M-X	0.51	0.13	15	.609

Note: Z = Wilcoxon signed rank z-statistic, r = effect size, n = sample size, p = two-tailed p-value. *$p < .013$ (Bonferroni correction applied).

Table 5.11: Linear mixed effects modelling (LMM) estimates of effects of LuminUs feedback (L-persec) upon mean information content (IC) and the number of musical change points CPs.

Dependent Variable	**Estimates of Fixed Effects**				
	Estimate	Std. Error	df	t	p
Gaze feedback:					
Mean IC (P1)	0.075513	0.260437	46.35	0.290	.773
Mean IC (P2)	−0.048567	0.226549	36.44	−0.214	.831
No. of CPs (P1)	0.387729	0.788520	24.46	0.492	.627
No. of CPs (P2)	0.396658	1.416513	42.39	0.280	.781
Motion feedback:					
Mean IC (P1)	−0.925560	0.281384	47.44	**−3.289****	.002
Mean IC (P2)	0.017931	0.133684	39.04	0.134	.894
No. of CPs (P1)	−1.880072	1.090359	50.094	−1.724	.091
No. of CPs (P2)	−0.157630	0.530703	24.505	−0.297	.769

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. Significance estimates: $^{**}p < .01$. All estimates were calculated using *variance components* covariance type for random effects and *diagonal* covariance type for repeated measures.

Table 5.12: Linear mixed effects modelling (LMM) estimates of effects of the number of glances and body motion (b-motion) upon the mean information content (IC). Results are provided for percussionists and pianists separately, and in each case, both the participants' own values, and those of their co-performer (Co) are used.

Parameter	**Estimates of Fixed Effects**				
	Estimate	Std. Error	df	t	p
Percussionists:					
No. of glances	0.010117	0.009140	5.70	1.107	.313
No. of glances (Co)	0.010560	0.008647	117.64	1.221	.224
Mean b-motion	−0.589507	0.207402	17.17	**−2.842***	.011
Mean b-motion (Co)	−0.674074	0.374785	12.93	−1.799†	.095
Pianists:					
No. of glances	0.003155	0.003735	132.69	0.845	.400
No. of glances (Co)	0.007359	0.002494	59.36	**2.951****	.005
Mean b-motion	0.190134	0.100006	8.13	1.901	.093
Mean b-motion (Co)	0.029211	0.066407	126.57	0.440	.661

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. Significance estimates: $^{*}p < .05$, $^{**}p < .01$. Estimates were calculated using *variance components* or †*unstructured* covariance type for random effects and *diagonal* covariance type for repeated measures.

number of glances upon the mean IC, using the data from all experimental conditions. The results are provided in Table 5.12. Again, the analyses were performed separately for pianists and percussionists. In each case, relationships were analysed between the participant's mean IC and i) their own values for mean body motion and glance count; ii) the mean body motion and glance count of their co-performer. The results show that for the percussionists there is a significant negative relationship between their own mean body motion and the mean IC ($t = -2.84$, $p = .011$). This relates to the result for motion feedback and mean IC for percussionists in Table 5.11, since it has already been shown that the body motions of each dyad are correlated (Table 5.9). Considered together, these results suggest a general negative relationship between body motion and the percussionists' mean IC. A possible explanation for this is that the IC values are greater when the percussionists stop and start more frequently, since the IDyOM model used to estimate the IC takes the duration between beats as its input (see Section 5.3.4). More frequent stopping would result in a lower average value for body motion.

Looking at the results for pianists, there is a significant relationship between the pianists' mean IC and the number of glances made by their co-performers (percussionists). This is an interesting result, since it provides quantitative and empirical support for the idea that gaze is used to send and receive information during musical interactions, as proposed by Schutz (1976), Davidson and Good (2002), and others. With regard to causality, there are a couple of possible explanations for this relationship: i) the glances of the percussionist acted as a stimulus for the pianist to change what they were playing; and ii) during periods of change (high IC), the percussionist looked towards the pianist for cues. It is quite possible that the observed effects were actually the result of a combination of these two explanations. Further work is required to support and investigate the nature of these findings.

5.4.4 Self-report Measures

This section concerns the analysis of the influence of the LuminUs feedback upon self-reported aspects of the performances, as recorded in the post-performance questionnaires. Initial analyses investigated whether the ratings differed between the conditions where both participants were receiving the same type of feedback: either gaze, motion, or no feedback. Figure 5.10 shows a stacked bar graph of the cumulative responses for each item in each of these three conditions. For items that could be considered to index positive attributes of the performance (1-4, 7, 8), the proportion of 'agree' responses relative to 'disagree' responses are higher in the LuminUs feedback conditions, than for the no-feedback condition (X-X). For items 5 and 6, which address *being ignored* and *leadership* respectively, there is little difference between the conditions. For item 9, regarding *boredom*, the proportion of 'moderately' and 'strongly disagree' responses is slightly greater in both of the feedback conditions (G-G and M-M). Friedman tests were conducted to test whether the differences in self-report responses across the seven conditions were statistically significant. Tests were performed for each of the SR items. The results are shown in Table 5.13, indicating that there were no significant differences for any of the items.

The musicians were also asked to rate the creativity (on a scale of 0-10) of each attempt within each condition. Figure 5.11 shows the mean creativity ratings for the percussionists and pianists for each attempt within each condition. Comparing the feedback and no-feedback conditions, there do not appear to be any trends in the creativity ratings. The graph indicates that the second attempts were generally rated more highly than the first. There was not any statistically significant difference in self-reported creativity depending upon the feedback condition ($\chi^2(6) = 4.475$, $p = .613$).

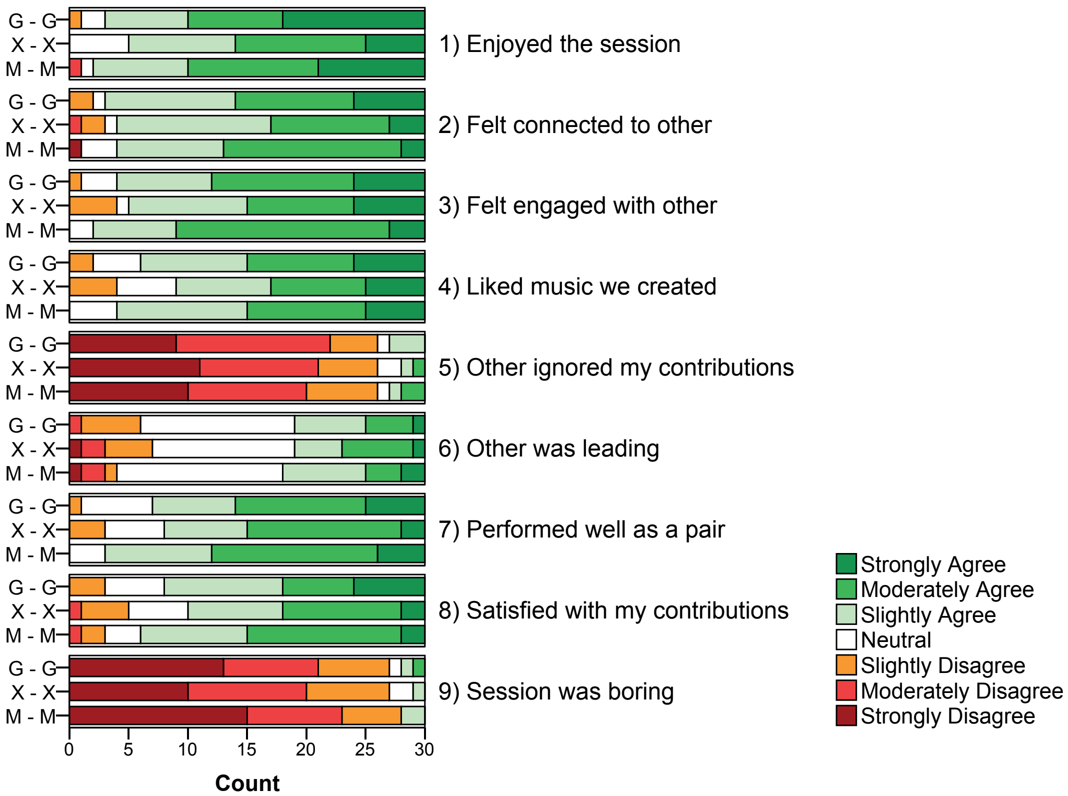


Figure 5.10: Stacked bar graph showing the cumulative self-report responses within the conditions where both participants receiving gaze feedback (G-G); motion feedback (M-M); or no feedback (X-X).

Table 5.13: Statistical results for differences between self-report responses across the seven LuminUs feedback conditions.

Self Report Item	χ^2	df	p
Enjoyed the session	4.673	6	.586
Felt connected to other	0.852	6	.991
Felt engaged with other	3.035	6	.804
Liked music we created	3.114	6	.794
Other ignored my contributions	1.168	6	.978
Other was leading	5.607	6	.469
Performed well as pair	4.577	6	.599
Satisfied with my contribution	3.059	6	.801
Session was boring	6.974	6	.323

Note: χ^2 = Friedman test statistic, df = degrees of freedom, p = significance level.

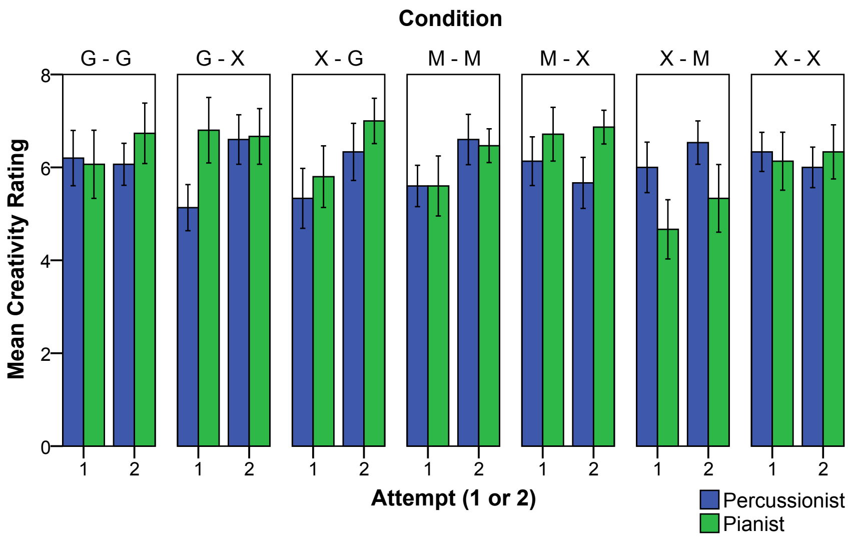


Figure 5.11: Bar graph showing the mean creativity ratings for the percussionists and pianists for each attempt within each of the conditions. G = gaze feedback, M = motion feedback, X = no feedback. Error bars: ± 1 SE.

5.5 Analyses and Results: Cardiac Activity and Musical Decision Making

This section analyses how the musical decisions made by the participants during their performances relate to various features of their cardiac activity. In order to perform these analyses, the following decision making features are used: musical change points (CPs); information content (IC); and animation content features (see Section 5.3.4). The participants' own self-reported creativity ratings were also used. Cardiac features were selected based upon those used in related studies of musical performance and cognitive tasks (see Section 4.2.3). Their extraction is described in Section 5.3.2. This section begins with analyses of potential correlations between discrete (time averaged) cardiac features and discrete creativity and decision making features. Subsequently, more detailed analyses of continuous-time features are undertaken.

5.5.1 Time-averaged Analyses

LMMs were used in order to analyse relationships between decision making and cardiac features that had been averaged over each accompaniment. Each of the analyses was performed separately for each type of musician: percussionist (P1) or pianist (P2); as well as for the entire group (All). Table 5.14 gives the results for the analyses of the effects of self-reported creativity upon the selected cardiac activity features. There are only a couple of noteworthy results. The first is the result for *SD HR* for percussionists (P1), which indicates that variation in HR is positively correlated with self-reported creativity ($t(14) = 2.31$, $p = .037$). The second result of interest is a negative relationship between the HR extrema count and self-reported creativity for pianists ($t(139) = -2.41$, $p = .017$).

Table 5.15 shows the equivalent analyses, using the mean number of musical CPs as a measure of musical decision making. There are no statistically significant results. However, it is worth noting that the three features that relate to variations in HR (SD HR, HR Extrema, and SDNN) are all negatively correlated with the number of CPs. Table 5.16 concerns relationships between IC and cardiac features. The results show that for the percussionists only, and for all participants combined, there are significant negative correlations between mean IC and the standard deviation in HR (SD HR) ($t(30) = -2.06$, $p = .049$; $t(47) = -2.17$, $p = .035$). For pianists only, there is a significant positive relationship between the number of HR extrema and mean IC ($t(93) = 2.25$, $p = .027$). For pianists and all participants combined, there are two significant positive correlations between mean IC and the frequency domain LF/HF

Table 5.14: Linear mixed effects modelling (LMM) estimates of the effects of self-reported creativity upon various cardiac activity features. Analyses are performed for percussionists only (P1), pianists only (P2), and all participants grouped (All).

Dependent Variable	**Par.**	**Estimates of Fixed Effects**				
		Estimate	Std. Error	df	t	p
Mean IBI	P1	−1.330	1.676	9.57	−0.794	.447
-	P2	−0.552	0.769	6.12	−0.719	.499
-	All	−1.278	0.858	12.61	−1.491	.161
SD HR	P1	0.100	0.043	14.11	**2.305***	.037
-	P2	−0.033	0.036	136.98	−0.930	.354
-	All	0.048	0.029	336.54	1.667	.096
HR Extrema	P1	0.015	0.174	24.30	0.087	.931
-	P2	−0.438	0.182	138.89	**−2.408***	.017
-	All	−0.153	0.146	69.99	−1.051	.297
SDNN (HRV)	P1	0.464	0.268	68.46	1.730	.088
-	P2	−0.230	0.276	136.28	−0.833	.406
-	All	0.244	0.213	339.20	1.148	.252
LF/HF ratio	P1	0.006	0.107	15.71	0.055	.957
-	P2	−0.157	0.111	20.34	−1.421	.171
-	All	−0.045	0.074	46.77	−0.607	.547
LF power	P1	7.444	8.594	67.89	0.866	.389
-	P2	−2.902	7.801	70.14	−0.372	.711
-	All	1.921	6.924	282.13	0.277	.782
HF power	P1	0.503	3.600	82.97	0.140	.889
-	P2	0.882	3.683	13.97	0.240	.814
-	All	0.428	2.752	210.96	0.155	.877

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. Significance estimates: $^{*}p < .05$. All estimates were calculated using *variance components* covariance type for random effects and *diagonal* covariance type for repeated measures.

Table 5.15: Linear mixed effects modelling (LMM) estimates of the effects of the number of change points (CPs) upon various cardiac activity features. Analyses are performed for percussionists only (P1), pianists only (P2), and all participants grouped (All).

Dependent Variable	**Par.**	**Estimates of Fixed Effects**				
		Estimate	Std. Error	df	t	p
Mean IBI	P1	1.593	0.909	12.83	1.752	.104
-	P2	0.588	0.907	13.94	0.648	.527
-	All	0.903	0.621	26.41	1.455	.158
SD HR	P1	−0.029	0.025	17.37	−1.126	.275
-	P2	−0.028	0.026	102.49	−1.099	.274
-	All	−0.011	0.020	322.69	−0.577	.565
HR Extrema	P1	−0.076	0.105	120.75	−0.725	.470
-	P2	−0.091	0.136	50.81	−0.669	.507
-	All	−0.079	0.091	304.86	−0.865	.388
SDNN (HRV)	P1	−0.046	0.172	54.92	−0.264	.793
-	P2	−0.232	0.209	150.57	−1.112	.268
-	All	−0.046	0.148	328.34	−0.308	.758
LF/HF ratio	P1	0.110	0.077	33.98	1.421	.164
-	P2	0.019	0.068	24.48	0.278	.783
-	All	0.064	0.049	33.34	1.312	.199
LF power	P1	2.081	7.629	18.73	0.273	.788
-	P2	2.081	7.629	18.73	0.273	.788
-	All	0.479	5.088	47.58	0.094	.925
HF power	P1	4.064	3.113	17.31	1.306	.209
-	P2	1.729	2.546	39.00	0.679	.501
-	All	3.713	2.290	43.00	1.622	.112

Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. All estimates were calculated using *variance components* covariance type for random effects and *diagonal* covariance type for repeated measures.

ratio feature of cardiac activity ($t(72) = 2.39$, $p = .020$; $t(98) = 1.99$, $p = .050$). Despite falling slightly short of statistical significance, it is also worth noting the result for SDNN (HRV) and all participants' data combined ($t(35) = -1.88$, $p = .068$)

Based upon the results presented in Tables 5.14, 5.15, and 5.16, four cardiac features show potential relevance to musical decision making: SD HR, HR extrema, SDNN, and LF/HF ratio. An additional observation was that the significant and noteworthy correlations for SD HR, HR extrema, and SDNN are of opposite sign in the IC analyses (Table 5.16) relative to those obtained for self-reported creativity (Table 5.14). This suggested that there may be a negative relationship between self-reported creativity and IC. To test this a LMM analysis was performed on the effects of IC upon self-reported creativity. The results indicated that mean IC is significantly negatively correlated with self-reported creativity ($t(48) = -2.10$, $p = .041$).

Table 5.16: Linear mixed effects modelling (LMM) estimates of the effects of information content (IC) upon various cardiac activity features. Analyses are performed for percussionists only (P1), pianists only (P2), and all participants grouped (All).

Dependent Variable	**Par.**	**Estimates of Fixed Effects**				
		Estimate	Std. Error	df	t	p
Mean IBI	P1	5.111	2.853	14.93	1.792	.093
-	P2	−3.062	3.527	11.88	−0.868	.402
-	All	2.234	2.450	29.79	0.912	.369
SD HR	P1	−0.172	0.083	29.63	**−2.058***	.049
-	P2	−0.138	0.157	111.05	−0.882	.380
-	All	−0.162	0.074	47.20	**−2.171***	.035
HR Extrema	P1	−0.195	0.388	63.52	−0.503	.617
-	P2	1.791	0.796	93.23	**2.249***	.027
-	All	0.145	0.356	331.66	0.407	.684
SDNN (HRV)	P1	−0.625	0.561	117.10	−1.114	.268
-	P2	−2.145	1.226	92.41	−1.750	.083
-	All	−1.069	0.568	35.32	−1.883	.068
LF/HF ratio	P1	0.248	0.199	41.57	1.246	.220
-	P2	0.885	0.371	72.48	**2.388***	.020
-	All	0.392	0.197	98.34	**1.987***	.050
LF power	P1	−16.046	16.540	84.78	−0.970	.335
-	P2	−30.348	35.531	78.18	−0.854	.396
-	All	−23.576	17.096	234.06	−1.379	.169
HF power	P1	2.445	11.338	14.62	0.216	.832
-	P2	1.478	15.594	48.39	0.095	.925
-	All	−2.055	7.016	192.95	−0.293	.770

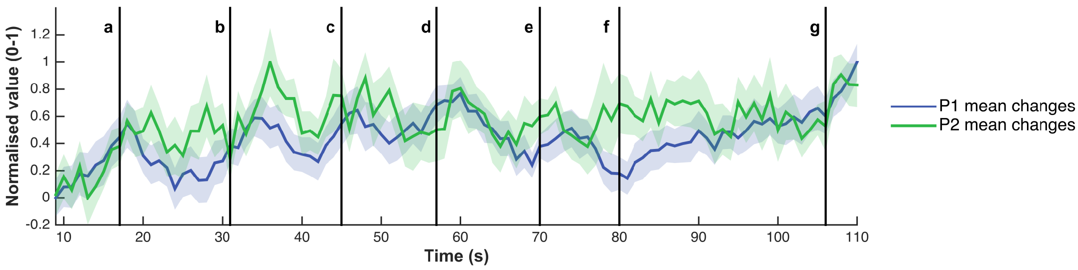
Note: t = t-value, df = degrees of freedom, p = two-tailed p-value. Significance estimates: $^{*}p < .05$. All estimates were calculated using *variance components* covariance type for random effects and *diagonal* covariance type for repeated measures.

5.5.2 Time Series Analyses

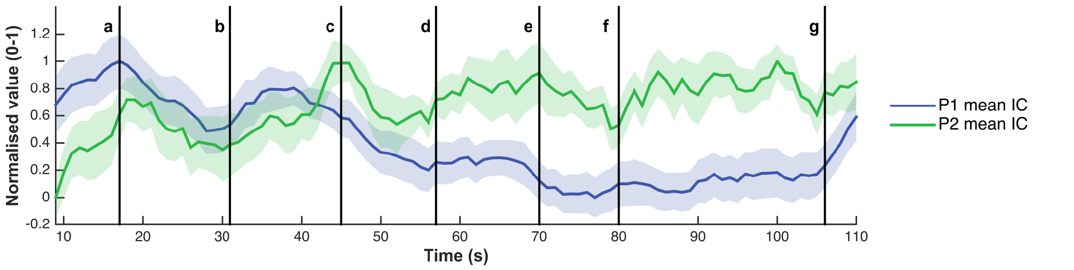
The purpose of the time series analyses was to investigate temporal relationships between cardiac features and musical decision making features. There were 14 time series for each participant (7 conditions $\times$ 2 attempts), resulting in a total of 420 time series for each feature (30 participants $\times$ 14 series). Consequently, the decision was made to commence the analyses by looking at averaged time series, facilitating the visualisation of general trends in the data. Each participant's time series feature was initially smoothed using a five second (5 sample) sliding average filter. This was done due to timing accuracy issues with the MIDI data (see Section 5.3.4 for further details). Subsequently, each participant's set of 14 smoothed time series were averaged; resulting in a single average series for each participant[4]. To account for variation between individuals, a unity-based normalisation (feature scaling) was applied to each of these series before averaging over each set of participants: percussionists (P1) and pianists (P2). Finally, an additional normalisation step was performed to facilitate comparisons between different features and instrumentalists. Animation cue point markers (labelled from **a** to **g**) are overlaid on each plot. These represent the time points in the animation where there are noteworthy changes in the content/storyline. These 'animation cues' are described in more detail in Section 5.3.4.

Figure 5.12 shows time series plots for the two MIDI-derived decision making features: the mean number of musical changes (derived from the CP feature) (see Fig. 5.12(a)) and the information content (IC) (see Fig. 5.12(b)). The musical changes time series for each accompaniment consisted of 120 values representing each second of the two minute attempt duration. These values were set to one at the time points where changes occurred, and zero otherwise. Figure 5.12(a) indicates that the musical change plots for percussionists and pianists share some similarities, especially during the first half of the performance. The average musical change series for the pianists appears to fluctuate more frequently than for the percussionists. This might be expected, since the piano had more degrees of freedom for musical change, consisting of 61 notes compared to the 4 drum pads provided to the percussionists. The time series for the percussionists (P1) shows a series of five peaks in the first 80 seconds. The rising edges of these five peaks coincide with the animation markers **a**, **b**, **c**, **d**, and **e** respectively. In each case the peak in the number of changes occurs less than 5 seconds after the marker. Between markers **f** and **g** there is a steady rise in the average number of change points, and following **g** this immediately begins to rise more steeply. These observations sug-

[4]In the case of the CP series, the smoothing was performed after averaging due to the binary (change/no-change) nature of the original series.



(a)



(b)

Figure 5.12: (a) Musical changes; and (b) information content (IC) averaged over all performances for percussionists (P1) and pianists (P2). Values have been normalised (0-1). Shaded area indicates ± 1 SE.

gest that the percussionists' playing was guided to some extent by cues in the content of the animation. For the pianists there is not such a clear relationship between the animation cues and the average musical change time series. Increasing changes appear to coincide with, or occur shortly after markers **a**, **b**, **d**, and **e**. Between markers **f** and **g** there are fewer large peaks in the pianists' changes and instead it appears to level off, with small variations.

Figure 5.12(b) shows the average IC series for percussionists and pianists. For the percussionists there is a generally decreasing trend in the IC between markers **a** and **f**. This suggests that the percussionists were varying their playing more towards the start of the animation. Between markers **f** and **g** the series remains fairly flat, before increasing steeply after the final animation cue. For the pianists there are two distinct peaks in the first 50 s of the time series, which appear to coincide roughly with markers **a** and **c**. The time series then roughly levels off, with two prominent dips occurring around markers **f** and **g**.

In summary, it appears that the two average musical change series show some similarity with each other, as well as some variations that coincide with animation cues. In comparison, the two average IC time series appear to vary more slowly over time, and show less correspondence to each other or the animation cues. In both

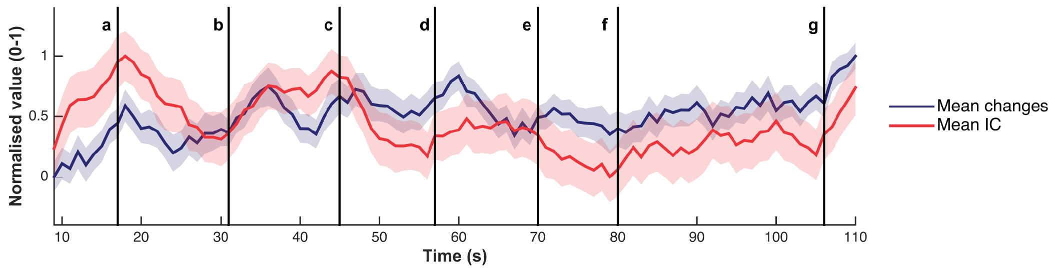


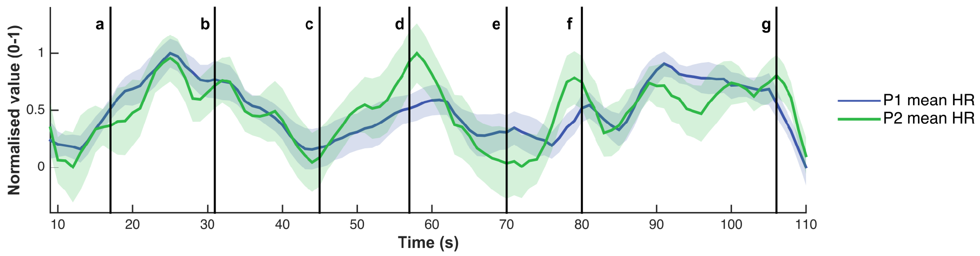
Figure 5.13: Musical changes and information content (IC) averaged over all performances. Values have been normalised (0-1). Shaded area indicates ± 1 SE.

plots the range of variation in the time series is greater during the first half of the accompaniment.

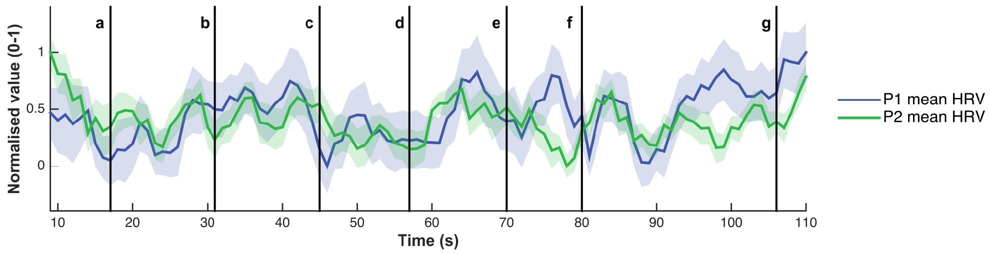
Figure 5.13 compares the two decision making features in a single plot. In this case the time series have been averaged over all participants. Once again, it appears that the range of variation for both averaged series is greater during the first half of the accompaniment. These variations appear to share some similarities between each other, with both series rising and falling in tandem for the majority of the accompaniment duration.

Figure 5.14 shows plots of the three continuous cardiac features that were chosen for the time series analyses. In Fig. 5.14(a) the average HR series for percussionists and pianists are plotted. The average HR series for each participant were initially de-trended, since most participants showed a generally increasing trend in HR over the course of each accompaniment, which could be associated with physical exertion. The two series show a high degree of similarity, especially during the first 40 s of the accompaniment, where there is a steady rise and drop in HR, peaking at around 25 s. Further peaks at around 58 s, 80 s, and 90 s also show a close correspondence between the two series. Regarding relationships between the two series and the animation cues, there are some interesting observations for the pianists' average time series. Specifically, it appears that animation cues **c**, **d**, **e**, **f**, and **g** are all closely concurrent with local HR extrema (peaks or troughs).

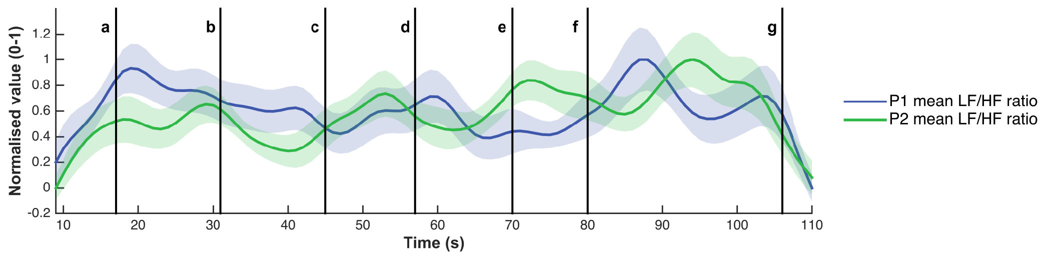
Figure 5.14(b) shows the heart rate variability (HRV) time series for percussionists and pianists. HRV is calculated as the first derivative of HR. Therefore, similarities between the plots for the two instrumentalists would be expected, as observed in the HR plot above. Indeed, both series share a similar pattern of peaks, predominantly in the first half of the accompaniment. Regarding visible relationships to the animation cues, it appears that the HRV tends to be falling, or within a local minimum (trough) at each of the cue points. Consequently, predominant peaks in HRV tend to occur



(a)



(b)



(c)

Figure 5.14: (a) Detrended heart rate (HR); (b) heart rate variability (HRV); and (c) LF/HF ratio averaged over all performances for percussionists (P1) and pianists (P2). Values have been normalised (0-1). Shaded area indicates ± 1 SE. Detrending was performed by subtraction of a 3rd order polynomial.

between the animation cues. This is especially apparent for the percussionists' time series between markers **c-d**, **d-e**, **e-f**, and **f-g**.

The final cardiac time series feature is the LF/HF ratio, which was calculated by performing spectral analysis upon sliding 30 s windows of the HRV series, with a 15 s overlap (see Section 5.3.2). As a result of this sliding window method, the series are smoother in appearance than the time domain cardiac features. For the percussionists' series, the dominant peaks in the series occur at around 20 s and 88 s, with smaller peaks at around 60 s and 105 s. These peaks do not appear to coincide with animation cues. For the pianists' series, there are four peaks, which are relatively evenly spaced. Again, these do not appear to share any relationship to the animation cue markers. There does not appear to be any noteworthy similarity between the P1 and P2 series.

Having performed a preliminary visual analysis of the time series features, time series analysis (TSA) techniques were adopted to investigate hypothesised relationships between musical decision making and cardiac activity. Numerous methods exist for analysing similarities between pairs of time series. These include dynamic time warping, spectral coherence, ARIMA modelling, and cross-correlation analyses. Cross-correlation methods were selected for the TSA in this study. Cross-correlation provides a measure of the covariance between two time series at varying time lags. This is particularly applicable to the analysis of physiological and behavioural data, where time differences often exist between physiological processes and observed behaviours. For a lag k, the cross-covariance between two time series, x and y is:

$$c_{xy}(k) = \begin{cases} \frac{1}{T}\sum_{t=1}^{T-k} (x_t - \bar{x})(y_{t+k} - \bar{y})\,; & k = 0, 1, 2, ... \\ \\ \frac{1}{T}\sum_{t=1}^{T+k} (y_t - \bar{y})(x_{t-k} - \bar{x})\,; & k = 0, -1, -2, ... \end{cases} \tag{5.5}$$

Where $\bar{x}$ and $\bar{y}$ are the sample means of the two series. The cross-correlation is then given by:

$$r_{xy}(k) = \frac{c_{xy}(k)}{s_x s_y} \tag{5.6}$$

Where s_x and s_y are the sample standard deviations of the two series.

An issue with human behavioural and physiological time series is that they often exhibit autocorrelation (correlation with themselves) (Dean and Dunsmuir, 2015). The individual time samples of an autocorrelated series are closely related to preceding values and cannot be treated as independent samples. Ignoring autocorrelation when performing cross-correlation analyses between time series can lead to the reporting of spurious correlations. Dean and Dunsmuir (2015) provide a comprehensive discussion of issues associated with performing cross-correlation on autocorrelated time series, with a specific focus upon human-derived time series, such as perceptual, performance, and movement data . To avoid these issues they suggest a general approach to performing TSA on a pair of autocorrelated time series. This approach involves two initial steps:

1. **Ensuring the series are stationary:** A stationary series is one with constant mean and variance (i.e. it does not have any trend). 'Stationarity' is commonly achieved by *differencing*, which involves creating a new series that consists of the differences between adjacent values of the original series.

2. **Removing autocorrelation:** Autocorrelation can be removed through a process known as 'prewhitening', which involves fitting an autoregressive (AR) model to one of the series and then using the parameters of that AR model to filter the same series. This has the effect of decorrelating the original time series, resulting in a series of uncorrelated residuals - a white noise series. The same filter is then applied to the second series.

In addition to removing autocorrelation using prewhitening, Dean and Dunsmuir (2015) also describe a method for estimating more realistic significance limits for the cross-correlation of the original autocorrelated series. The standard 5% significance limit $\pm L$ for cross correlation is given by:

$$L = 1.96/\sqrt{n} \tag{5.7}$$

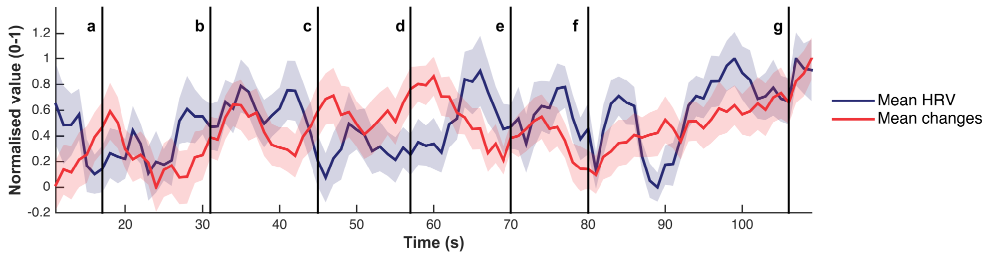
Where n is the sample size of the time series. The corrected significance estimate L_{corr} used by Dean and Dunsmuir (2015) and described by Cryer and Chan (2008), is given by:

$$L_{corr} = 1.96\sqrt{\frac{1}{n}\left[1 + 2\sum_{k=1}^{\infty} \rho_k(x)\rho_k(y)\right]} \tag{5.8}$$

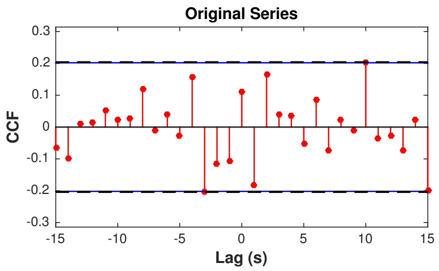
Where $\rho_k(x)$ and $\rho_k(y)$ are the autocorrelations at lag k of series x and y respectively.

In the following analyses each time series is initially tested for stationarity using the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. If either series is not stationary then differencing is performed on both series, followed by a further KPSS test for stationarity. The cross-correlation function (CCF) is then computed, using lags of $\pm k = 0, 1, 2...15$. Results are plotted along with the corrected significance limit estimates. Following this, prewhitening is performed on the stationary series and then the post-prewhitening CCF is calculated and plotted. The analyses were performed using MATLAB scripts, which were based upon the R software examples provided by Dean and Dunsmuir (2015).

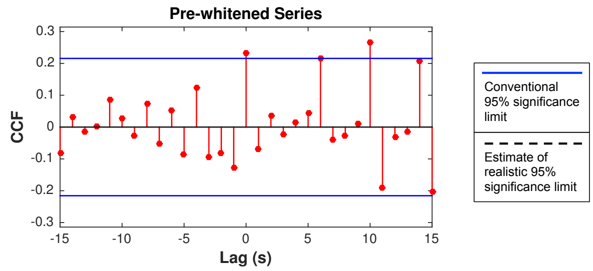
Figure 5.15 shows the results of the TSA for the mean HRV and the average musical change series from all percussionists. A visual inspection of Fig. 5.15(a) indicates that there may be some similarities between the two series at certain points. The CCF of the original series, shown in Fig. 5.15(b), does not show any significant correlations across all lags. However, the CCF of the prewhitened series suggests that there are significant positive correlations at lags of zero and 10 s. A positive lag signifies that the HRV series correlates with the change series after the former has been shifted forward in time (to the right), in this case by 10 s.



(a)



(b)



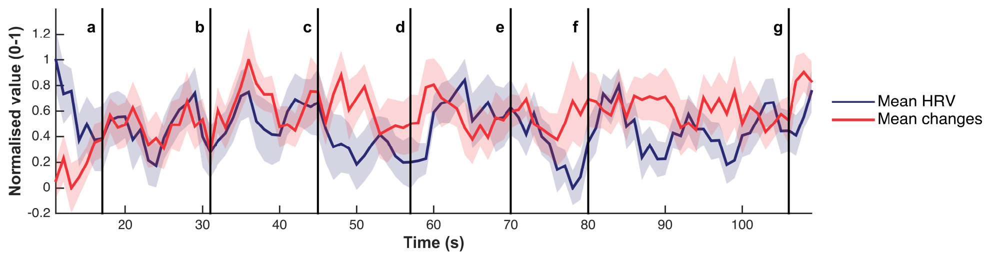
(c)

Figure 5.15: Analysis of time series correlations between musical changes and HRV over all performances for percussionists (P1) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating ± 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.

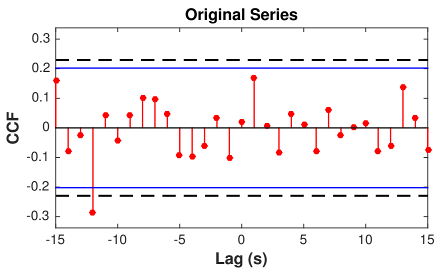
The equivalent results for pianists are shown in Fig. 5.16. Again, a visual inspection of Fig. 5.16(a) indicates that there may be some similarities between the two series, especially during the first 40 s. However, the CCF plots, shown in figures 5.16(b) and 5.16(c), suggest contrasting negative and positive correlations at lags of -12 s and -11 s respectively. The fact that the dominant correlations are of different signs, and are only marginally greater than the significance limits, suggests that there is not a causal correlation between the two series.

Figure 5.17 shows the results of the TSA for the mean HRV and mean IC series from all percussionists. A visual inspection of Fig. 5.17(a) immediately suggests that there is not any correlation between the two series. This is confirmed in figures 5.17(b) and 5.17(c), where no significant correlations exist across all the time lags.

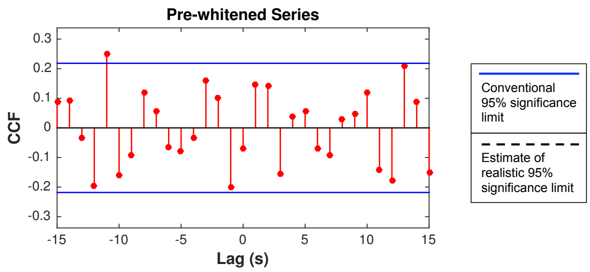
The equivalent results for pianists are shown in Fig. 5.18, and a visual inspection of Fig. 5.18(a) indicates more similarity between the series than was observed for the percussionists. However, the CCF results do not indicate that these series are correlated. Two significant negative correlations are found at lags of ± 13 s for the original series (5.18(b)), but they are undermined by the fact that they are far apart, only marginally significant, and not reproduced in the prewhitened CCF (5.18(c)).



(a)

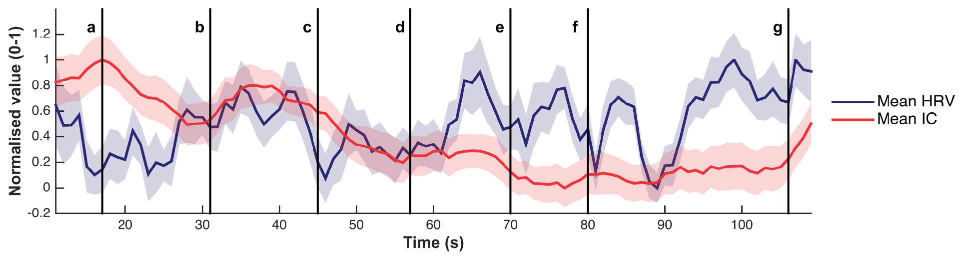


(b)

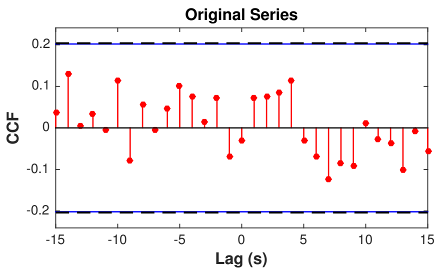


(c)

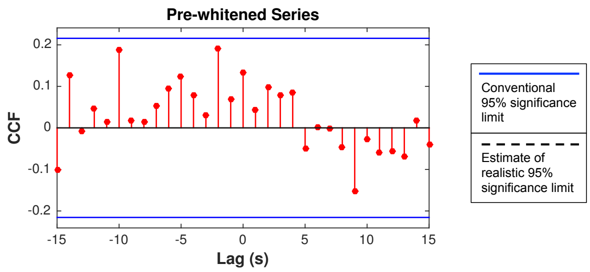
Figure 5.16: Analysis of time series correlations between musical changes and HRV over all performances for pianists (P2) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating ± 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.



(a)



(b)



(c)

Figure 5.17: Analysis of time series correlations between mean information content (IC) and HRV over all performances for percussionists (P1) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating $\pm$ 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.

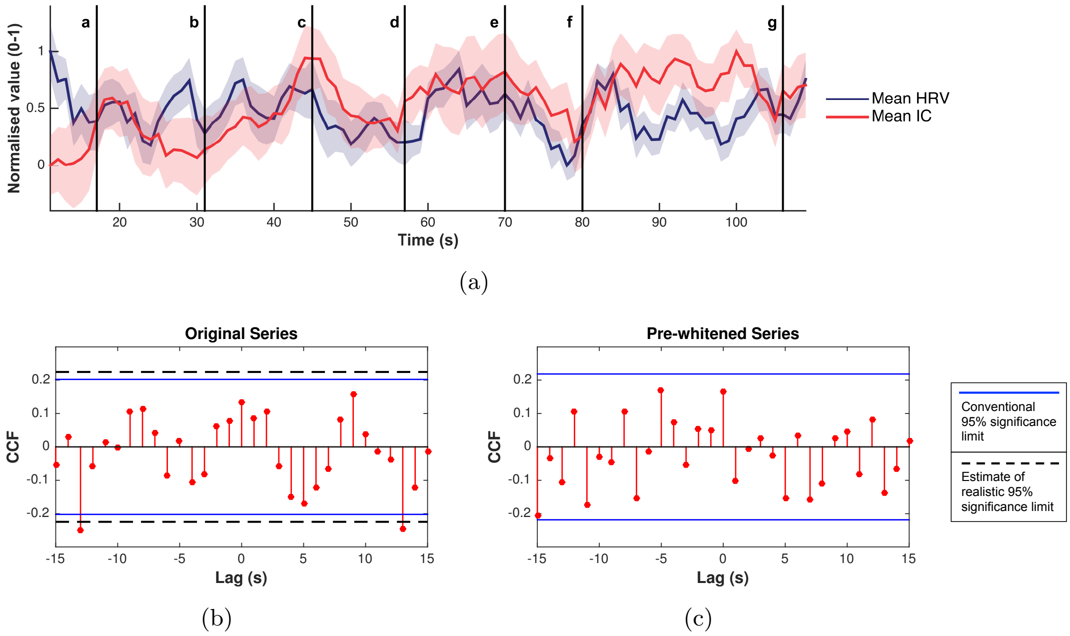


Figure 5.18: Analysis of time series correlations between mean information content (IC) and HRV over all performances for pianists (P2) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating ± 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.

Equivalent analyses to those for HRV were performed; but using LF/HF ratio as the cardiac time series feature (see figures B.4.1 to B.4.4 in Appendix B.4). In each case there were no noteworthy correlations. One of the issues is that the LF/HF ratio series exhibit a more gradual variation, which leads to a high degree of autocorrelation. Consequently, the prewhitened time series still exhibit autocorrelation, as indicated by the smooth appearance of the CCFs.

The analyses above use time series that have been generated by averaging across the individual time series of multiple participants. This is a valid approach for the investigation of relationships between cardiac activity and musical decision making, since the process of averaging serves to remove random factors that influence cardiac activity during individual accompaniments. However, an observation made during Study 1 was that apparent links between cardiac activity and rhythmic change points were participant-specific. For a more individual-focused analysis, the cross-correlations of averaged time series were computed for individual participants. The processing of the time series was identical to above: first, testing for stationarity and applying differencing if necessary; then applying prewhitening to the series. Only correlations between decision making and HRV were analysed; since the group TSA indicated that the slow changing and highly autocorrelated nature of the LF/HF ratio series does not make them particularly suitable for cross-correlation analyses. Rather than plotting the results for all participants, the largest (absolute) cross-correlation value ($r_{xy}(k)$) and its associated lag time (k) were extracted from each participant's CCF. The results are shown in Table 5.17. In addition to the largest correlation value, the column labelled 'Dsig' provides the difference between the largest correlation value and the significance bounds for the CCF. Therefore, values of Dsig indicate the extent to which the correlation value is greater than (positive values), or less than (negative values) the significance value. Correlation values that surpass the significance bounds have been marked with an asterisk.

Looking at the results for the CCFs between mean HRV and the number of changes (Table 5.17 (a)), it can be seen that the majority of correlations for both percussionists and pianists are not significant. Furthermore, of those that are significant, the correlation and Dsig values are small (< 0.3 and < 0.1 respectively). Based upon these results, one would not conclude that HRV alone can serve as a predictor of musical change for individuals.

Table 5.17: Results of cross correlation analysis between HRV and (a) number of changes; and (b) information content (IC). The table presents the greatest (absolute) correlation, $r_{xy}(k)$, and associated lag, k, for each participant's (P1 - percussionists, P2 - pianists) averaged and prewhitened time series. It also includes *Dsig* - the amount by which the correlation value exceeds (positive values) or falls short (negative values) of the CCF significance bounds.

Dyad	**(a) Number of changes**			**(b) Information content**		
	$r_{xy}(k)$	k	Dsig	$r_{xy}(k)$	k	Dsig
Percussionists:						
1	−0.191	1	−0.024	0.347*	−2	0.134
2	−0.162	6	−0.047	0.250*	14	0.042
3	−0.178	−5	−0.029	0.144	0	−0.063
4	0.187	−3	−0.026	−0.234*	13	0.017
5	−0.302*	11	0.078	−0.262*	4	0.040
6	−0.311*	−11	0.092	0.241*	−1	0.022
7	0.246*	−7	0.028	0.138	−4	−0.075
8	0.221	−9	−0.002	−0.226*	−3	0.004
9	−0.198	−10	−0.026	−0.282*	0	0.069
10	−0.252*	−3	0.028	0.247*	−12	0.023
11	0.161	−12	−0.046	0.231*	6	0.024
12	0.244*	−14	0.023	−0.231*	−3	0.011
13	0.172	−3	−0.041	0.195	1	−0.018
14	0.203	4	−0.021	0.297*	4	0.072
15	0.196	−14	−0.012	0.245*	−15	0.031
Pianists:						
1	0.254*	6	0.046	−0.312*	9	0.104
2	−0.180	−14	−0.045	0.162	−8	−0.063
3	0.111	−7	−0.113	−0.266*	4	0.045
4	−0.245*	−2	0.035	−0.325*	2	0.116
5	0.223*	0	0.007	−0.197	−6	−0.018
6	−0.228*	−8	0.014	−0.209	15	−0.005
7	−0.217	−3	−0.001	−0.251*	6	0.033
8	0.260*	13	0.042	−0.329*	−4	0.111
9	−0.163	−11	−0.053	−0.240*	−13	0.023
10	0.165	−1	−0.049	−0.205	4	−0.020
11	0.216*	15	0.007	0.295*	−2	0.081
12	−0.194	−7	−0.020	0.236*	−13	0.016
13	−0.203	13	−0.016	0.309*	−14	0.084
14	−0.210	8	−0.002	−0.280*	−5	0.068
15	0.233*	−14	0.011	0.253*	2	0.028

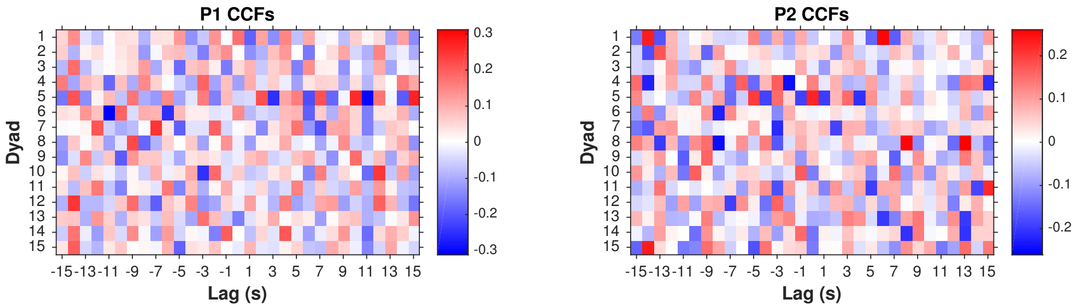
*Note: *Significant values of* $r_{xy}(k)$.

The results for information content (IC) (Table 5.17 (b)) show that the majority of correlations are significant. For percussionists, five of the significant correlations are negative and seven are positive; whereas, for pianists, seven out of the eleven significant correlations are negative. These findings suggest that there is unlikely to be a general relationship between HRV and IC that can be applied across musicians. However, the findings do not rule out the possibility of relationships that are specific to, and characteristic of individuals. Looking specifically at the results for pianists, it can be seen that the pianists in dyads 1, 4 and 8 all have significant negative correlations, with Dsig greater than 0.1. However, the lag values for each participant are different in both sign and quantity, suggesting substantial variation between individuals.

To gain further insight into the cross-correlations of HRV and decision making features for individuals, plots were created in order to visualise the correlations across all lag values for each participant. Each row of the plots represents an individual participant and each column represents a lag value (from -15 s to 15 s). The colours of the squares represent the value of the correlation for that particular participant at that particular lag. The plots for mean HRV and mean changes are shown for percussionists and pianists in figures 5.19(a) and 5.19(b) respectively. By presenting the results in this way, it is possible to look for signs of time series relationships at the levels of both the individual, and the group. For the former, rows of the plots should be inspected individually for the presence of sequences of correlations that are of the same sign (colour), and of high value (colour intensity). For the latter, the entire plot should be inspected to for the presence of vertical strips of the same colour that may indicate a common correlation and lag. In Fig. 5.19 the distribution of correlations in both plots does not appear to reveal any such patterns, either at the level of the individual, or the group.

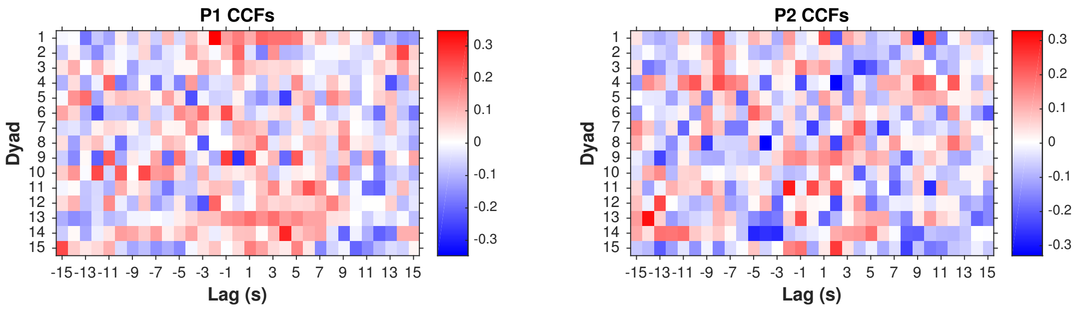
Figure 5.20 shows the equivalent plots for the cross-correlations between mean HRV and mean IC. Once again, there do not appear to be any immediately visible patterns, either at the level of the individual, or the group.

Coupled with the results for the group-averaged time series analyses, the results for the analyses of individual correlations suggest that there are no direct temporal correlations between cardiac features and musical decision making.



(a) (b)

Figure 5.19: Plot of the CCFs between the mean HRV and the mean number of changes time series for each participant. Each row represents the results for a single participant, and each column represents a lag value. The colour of each box represents the correlation value, where red values are positive and blue values are negative. All series were prewhitened prior to calculation of the CCF.



(a) (b)

Figure 5.20: Plot of the CCFs between the mean information content (IC) and the mean HRV time series for each participant (see Fig. 5.19 caption for a detailed description of the plot format).

5.6 Discussion

The study reported in this chapter was designed to test multiple exploratory questions and hypotheses. These were split according to two aims. The first aim concerned the effects of the LuminUs feedback upon the experiences and behaviours of the musicians. The second aim concerned the investigation of relationships between cardiac activity and musical decision making. Given that these aims were addressed separately during the analyses, distinct sections are dedicated to each in the discussion below. However, there are also aspects of the findings across the entire study that raise interesting points for general discussion. Accordingly, these points are addressed at the end of this section.

5.6.1 Effects of the LuminUs Feedback

The LuminUs was designed with simplicity and versatility in mind; as a device that could aid the non-verbal interactions between co-present performing musicians. It was not experimentally possible to test the LuminUs under the array of possible circumstances where it could be employed. Instead, this study aimed to investigate the general effects that it might have upon dyadic musical improvisation. The analyses were predominantly quantitative, focusing on four types of data collected during the experiments: i) gaze data; ii) motion data, iii) MIDI data; and iv) self-report data.

Gaze Behaviour

From the eye-tracking data it was possible to look at how the glancing behaviours of the musicians were influenced by the LuminUs. The results show that, when using the LuminUs, the musicians would make a greater number of glances than when not using it. This effect was observed regardless of the type of feedback being provided: gaze or motion. However, effect sizes were greater for the gaze feedback conditions.

The average duration of individual glances did not change in relation to the LuminUs feedback. This might be expected, given that once a glance is initiated, the musician is no longer under the influence of the LuminUs, since they are, by definition, looking away from it. Therefore, glance duration is more likely to be influenced by communicative and social factors associated with gaze.

To test whether the musicians' glancing behaviour was directly influenced by the gaze feedback, two approaches were used: firstly, by analysing the proportions of reciprocated glances; and secondly, by looking for correlations between quantitative measures of LuminUs feedback and the number of glances. For the former, the proportion of glance reciprocation (with a delay of 5 s) was found to be significantly higher in the condition where both musicians had gaze feedback, relative to the condition where they had none. For all the other conditions there were no effects. LMMs were then used to test whether quantitative measures of the LuminUs feedback – L-time and L-persec – had any effects upon the number of glances. Significant effects were found for both LuminUs measures, but only in the gaze feedback conditions. Considered together, these results suggest that the musicians were responding directly to the gaze feedback by glancing back towards their co-performer.

Throughout the gaze analyses, observed effects of the LuminUs were found to be greater for the percussionists than the pianists. There are a number of reasons why the effects of the LuminUs might be expected to vary according to the instrument being played. Firstly, the ability to observe the LuminUs feedback during a performance is likely to be influenced by the placement of the LuminUs in relation to the instrument. In this experiment the placement was identical for both the percussionists and pianists. However, the cymbal drum pad was positioned higher than the piano, and therefore, vertically closer to the LuminUs. Furthermore, the percussionists appeared to spend less time looking at their instrument than the pianists. These factors may have resulted in the percussionists having more opportunity to observe the LuminUs feedback, resulting in the greater effects observed for them. A second reason why the effects might be instrument dependent relates to the musical dynamic between the two musicians. For example, the pianists tended to glance towards the percussionists more than the percussionists glanced towards the pianists. This can be seen in figures 5.7(a) and 5.7(b). This may have been due to the fact that the pianists were looking for rhythmic cues, and that it is harder for the percussionists to take cues from the less visible movements of the pianists' fingers. Again, this would have resulted in the LuminUs feedback being more salient to the percussionists.

Body Motion

The analyses of the effects of the LuminUs feedback conditions upon body motion did not result in any significant effects. Subsequent analyses showed a significant correlation between the amount of motion feedback provided by the LuminUs (as measured by the L-persec feature) and the body motion of the participant receiving that feedback. However, this was attributed to the fact that the body motions of participants within dyads were found to be significantly correlated.

Musical Decision Making

The number of change points (CPs) and the mean information content (IC) features were used to investigate whether the LuminUs had any effects upon musical decision making. Analyses of the influence of the seven feedback conditions upon the musical decision making features were conducted for percussionists and pianists separately. No statistically significant differences were found for pianists, whilst the mean IC for percussionists was found to be significantly higher in the condition where both participants had motion feedback (M-M) relative to the no feedback condition. However, this result was not repeated for the other condition where percussionists had motion feedback, suggesting that there was no general influence of the LuminUs upon musical decision making.

Subsequent LMM analyses revealed that there was an underlying correlation between the pianists' mean IC and the number of glances made by their co-performers across all experimental conditions. This result provides support for the roles of gaze in conveying and gathering information during collaborative music making. Considering the latter role, it was suggested that a casual explanation for this result was that the percussionists glanced towards the pianists during periods of change (high IC) as a means of gathering information (or cues) about the nature of the musical change. Indeed, Sawyer (2003) notes that visual contact is an important part of anticipating and communicating change during improvised performance. On the other hand, considering gaze as a conveyor of information – a means of non-verbal communication – it was suggested that the glances of the percussionists served to initiate the pianists' musical changes. Once again, attention must be drawn to the fact that this result was instrument dependent. The fact that no comparable result was observed for the percussionists could be due to various factors. These factors could relate to differences between the instruments and their functional roles within the musical arrangement. For example, it is possible for the pianist to introduce both rhythmical and tonal changes (e.g. a change of key), whilst the drummer can only introduce rhythmic change. Consequently, the pianist could make a tonal change and it would not be necessary for the drummer to change their playing in any way. Whereas, for the drummer to make a rhythmic change (especially a change in tempo) it would be more important for them to signal their intent to the pianist prior to instigating the change, so that the pianist could change their playing accordingly.

It is also possible that the lack of a comparable result for percussionists was due to the difference in the way that the IC was calculated for each instrument. In particular, the IC for pianists was calculated using pitch intervals; whilst the IC for percussionists was calculated using beat onset timings. This means that the percussionists' IC values would have been more sensitive to breaks in their playing. This potentially explains why a significant negative correlation was found between the percussionists' mean IC and their body motion (see Table 5.12).

In summary, the analyses suggest that the LuminUs did not have any effect upon decision making. However, through the use of quantitative measures of gaze and musical decision making, it was possible to provide empirical support for the communicative roles of gaze in musical interaction. It is worth noting that the analyses do not consider any qualitative aspects of decision making and musical change. Such analyses would have required the musicians to manually review and comment upon their decisions, which was not feasible from a practical standpoint. The following discussion of the self-report results provides some insight into qualitative aspects of the LuminUs's influence.

Self-report Measures

The analyses of self-report (SR) ratings showed that none of the items differed significantly across the seven LuminUs feedback conditions. Despite not achieving statistical significance, all of the SR items relating to positive aspects of the performance received a higher proportion of agreement ratings (relative to disagreement) in the conditions where participants were receiving some form of feedback from the LuminUs. Lack of statistical power may have been a result of the small sample size and a degree of variability in the extent to which participants were influenced by the LuminUs feedback.

Looking at differences between the two types of feedback (motion and gaze), the proportion of agreement ratings for items relating to positive aspects of the performance were generally greater for the motion feedback condition. This difference may have been related to fact that the LuminUs was generally providing some level of light feedback throughout the motion feedback conditions; whereas for gaze feedback the light would only appear for comparatively short periods during glancing. These results pose some interesting questions relating to the way in which the two types of LuminUs feedback might have been interpreted. To what extent were the observed differences related to the meaning conveyed by the feedback? For example, participants agreed with the statement – 'I felt connected to the other musician' – more for the gaze feedback condition; whilst the item – 'I felt engaged with the other musician' – received more agreement ratings for the motion feedback condition. Is there a semantic distinction between 'connection' and 'engagement' that relates to the distinction between gaze and motion feedback? Given the predominantly quantitative nature of this study, it is difficult to provide any substantiated answers to these questions. However, they highlight interesting considerations that could be addressed in future work.

The SR questionnaire also asked participants to rate the creativity of each of the two accompaniments within each condition. The analyses did not reveal any effects of the LuminUs feedback upon self-reported creativity. This is, perhaps, unsurprising, since creativity is likely to involve a substantial amount of variability owing to the wide array of factors that can influence it. Furthermore, creativity is highly subjective and SR ratings may not be the most appropriate method for measuring it. Recordings of the accompaniments could have been used to collect creativity ratings from external judges. However, given the large number of accompaniments (210 in total), this would have been a time consuming task and was not within the scope of this study.

5.6.2 Cardiac Activity and Musical Decision Making

This section discusses the findings relating to the discrete and time series analyses of relationships between cardiac activity and musical decision making.

Time-averaged Analyses

For the analyses of relationships between creative decision making and cardiac activity, discrete, time-averaged data were initially used. The average number of change points (CPs) and average information content (IC) were the two discrete features that were used as measures of decision making. Again, it should be noted that neither of these features provide a qualitative measure of the decision making (e.g. how original the decisions were). In an attempt to address this, the self-reported creativity ratings were also used. Conversely, these ratings do not provide any measure of the quantity of musical decision making. However, it was thought that comparing the results across all three measures would provide a broader perspective on the potential relationships between musical decision making and cardiac activity.

For the creativity rating analyses the standard deviation in heart rate (HR) (SD HR) was found to be significantly positively correlated with creativity for percussionists; whilst the number of HR extrema was significantly negatively correlated with creativity for pianists. SD HR and the HR extrema count are both measures of variation in HR. The former provides a general measure of the amount of variation in HR, whilst the latter is more specifically a measure of the amount of cyclic variation (since it counts successive peaks and troughs in the HR time series). Drumming generally involves a large amount of physical exertion, in comparison to other instruments. Therefore, HR levels might be expected to be somewhat associated with playing activity. Indeed, a post-hoc analysis of the effects of the number of beats played in an accompaniment upon the average HR shows a significant correlation[5] ($t = 5.068$, $p = .000$). As such, the correlation between creativity and SD HR for percussionists may be a result of varied playing activity being associated with high creativity. Additionally, the short length of the accompaniments meant that the participants' HRs would tend to rise gradually over the course of the accompaniment, as observed in the time series analyses (see Section 5.5.2). Consequently, the correlation for percussionists could also be associated with the overall increase in their HR in relation to their general levels of physical and mental exertion. This explanation is supported by the results from Study 1, where the mean body motion of the percussionists was found to be positively correlated with their

[5]Using linear mixed modelling with *variance components* covariance type for random effects and *diagonal* covariance type for repeated measures.

self-reported creativity scores (see Section 3.5.2). Given that HR is likely to stabilise over longer periods of playing, this would also explain why comparable correlations between creativity and SD HR were not observed in Study 1.

Playing activity is also likely to have effects upon HR for pianists. However, these effects are likely to be different from those for the percussionists, due to the differing nature of the playing. Again, a cursory analysis of the effects of the number of played notes upon the mean HR for pianists shows a significant effect ($t = 2.953$, $p = .014$). However, this effect is less than that for the percussionists. Furthermore, a significant effect was found for the mean velocity of the played notes upon the pianists' mean HR ($t = 4.520$, $p = .007$). No equivalent effect was observed for the percussionists. These differences could explain why equivalent correlations with SD HR were not observed for pianists. Instead, a negative correlation was found between self-reported creativity and the number of HR extrema. This indicates that when the pianists' HR exhibited less cyclic change over the course of an accompaniment their creativity ratings were higher. Given the aforementioned correlations between playing and HR, this suggests that playing consistency could have had a factor in determining self-reported creativity.

No significant correlations were found between the number of change points (CPs) and cardiac features. Based upon the findings from Study 1, it was hypothesised that the number of HR extrema might be positively correlated with the number of CPs. A potential issue is the fact that an automated method was used to extract CPs from MIDI data in the current study; whereas CPs were manually labelled in the first study. The automatic method was necessary due to the size of the data set, but might have resulted in less accurate measures of change.

The analyses of correlations between information content (IC) and cardiac features revealed interesting results. For percussionists, a negative correlation was found between IC and SD HR. As previously noted, the IC values for percussionists are likely to be highly influenced by breaks in playing. The fact that SD HR will be higher when a participant's HR increases more steeply over the course of an accompaniment, has also been discussed. Consequently, a likely explanation for this result is that breaks in playing led to higher average IC, but also allowed the percussionists' HRs to return towards a resting state; resulting in a lower overall increase in HR over the course of an accompaniment.

For the pianists, a positive effect of IC upon the number HR extrema was found. This is an interesting result, since it supports the findings from Study 1, which associated decision making with HR extrema. Again, the absence of an equivalent result for percussionists could be due to the difference in the method used for calculating IC.

A positive effect of IC upon LF/HF ratio was found for both percussionists and pianists separately, and when grouped. This effect was significant for the pianist and grouped analyses. This finding supports the hypothesised positive relationship between LF/HF ratio and decision making. LF/HF is thought to reflect modulation of the sympathetic nervous system, distinguishing sympathetic effects from parasympathetic effects (Peifer, 2012; Kim and André, 2008). The fact that it is a frequency-domain feature, calculated using de-trended inter-beat intervals, means that it is less susceptible to influence from time domain trends in cardiac activity relating to physical exertion.

When comparing the results of the self-reported creativity analyses to those obtained for the IC analyses, it was noted that the signs of the observed correlations were opposing. A subsequent test showed a significant negative effect of IC upon creativity ratings. This interesting result suggests a distinction between creativity viewed from a critical, self appraisal stand point, versus creativity from a pure sense of creating something new (a musical change). In this case, greater levels of IC could be indicative of a large number of musical change decisions, leading to a 'fractured' performance; which, when viewed as a whole, ends up lacking coherence as a creative piece.

In summary, the discrete time averaged analyses revealed that there are potential relationships between measures of cardiac activity and measures of musical decision making. In the case of SD HR and the HR extrema count, these relationships were attributed to changes in physical exertion that would have accompanied changes in playing activity. As discussed, SD HR could have been partially representative of the overall increase in HR over the playing period, due to the short length of the accompaniments providing insufficient time for HR to stabilise. Consequently, SD HR might not be such an informative feature for analysis of longer playing periods.

The analyses in this study were especially motivated by an interest in how the mental processes involved in decision making might be reflected in cardiac activity. The LF/HF ratio appears to be a more promising feature for this type of analyses, given that it is less prone to influence from physical activity. For future work, it may also be worth considering ways in which the physical activity of the musicians could be accounted for. This could involve using a combination of measures of motion sensing and playing activity; the latter of which has already been shown to be related to HR.

The analyses in this study also raise interesting questions regarding the ways in which quantitative aspects of decision making might relate to overall evaluations of creativity. By asking musicians to elaborate on the factors that influence their self-reported ratings of creativity, future studies could shed some light on these questions.

Time Series Analyses

One of the motivations for investigating relationships between cardiac activity and musical decision making is that it could be incorporated into the LuminUs design as an extra feedback modality; providing real-time visualisations relating to the decision making intentions of improvising musicians. In order to achieve this, it would be necessary to model continuous relationships between musical decision making and cardiac activity. Consequently, the decision was made to perform a broad time series analysis (TSA) of the continuous cardiac and musical decision making data. In this case, only the two quantitative measures of decision making (CPs and IC) were used; since no continuous measures of self-reported creativity were collected. Animation cues were also considered as discrete indicators of time points where the musicians might have been stimulated to change their playing.

Due to the large data-set, the approach used in the present study was to commence by averaging over participants, and then to look at individual participant relationships for specific features of interest. Attempting to perform TSA that was sensitive to both individual and group-level relationships was one of the main challenges of the analyses. During averaging, unity-based normalisation was used to account for individual variability in the data. This is a valid approach, which is often used in cases where physiological signals are compared over groups (Mandryk et al., 2006; Chang et al., 2009). However, this normalisation method is also sensitive to outliers (Broek et al., 2009) and, therefore, has the potential to introduce unwanted errors into the TSA.

Visual inspection of musical decision making series: The initial visual analyses of the averaged CP time series showed distinct similarities between the percussionists' and pianists' series. Furthermore, rising edges of the series appeared to coincide with the animation cues; suggesting that the animation encouraged the participants to change their playing at certain points. This also provides a good indication that the automatic method for extracting CPs was capable of providing meaningful information about where musical changes occurred.

The IC series showed fewer similarities between percussionists and pianists. The percussionists' series appeared to gradually decrease over the first minute of the performance, indicating that they were changing, or taking more breaks in their playing during this phase. Indeed, observations of the performances indicated that it was more common for the pianist to start the accompaniment and the percussionist to join in at a later point. They would then tend to 'settle into a rhythm', which is reflected by the lower and less variable IC in the second half of the accompaniment. Again, both series appeared to show some relation to the animation cues; however, this was not as evident as for the CP series.

Similarities were noted when comparing the CP and IC time series. Both series generally showed a higher amount of fluctuation during the first half of the performance, before settling and then rising during the final 15 seconds. This is supported by the researcher's observed experience of the performances, whereby participants tended to vary their playing more at the start of the accompaniment and then attempt to conclude the accompaniment by changing their playing during the final seconds. Again, these observations are evidence that the two measures – CP and IC – provided meaningful representations of the musical decisions made during the performances.

Visual inspection of cardiac feature time series: Similar visual analyses were performed on the cardiac feature time series. The two detrended HR series for percussionists and pianists appeared to show distinct similarities, especially during the first half of the performance. Prior results in the present study showed that the average body motions of performers within dyads are significantly correlated; and that there are similarities in the CP time series for both instrumentalists. Consequently, a degree of synchrony would be expected between the cardiac activity of the performers. For the pianist time series, a number of alignments between HR extrema and the animation cue points were observed. This observation supports the findings for the discrete time averaged analyses, which showed a positive correlation between HR extrema and IC for pianists.

For the HRV time series, observations showed that, for both instrumentalists, HRV tended to be falling or within a local minimum (trough) at each of the animation cue points. This finding is supported by the results from the discrete analyses, which showed that SDNN (HRV) was negatively correlated with the mean CP count and mean IC. For IC these effects were close to significance for all participants ($p = .068$), and for pianists only ($p = .083$). Considered collectively, these findings support the hypothesis that HRV would be negatively correlated with the number of musical changes (see hypothesis **cH2**). This hypothesis was influenced by studies showing that HRV decreases with stress and mental workload (Mandryk et al., 2006; Rowe et al., 1998).

The third cardiac feature that was used in the TSA was LF/HF ratio. Despite being a promising feature for discrete time analyses (see discussion above), the fact that LF/HF ratio is a frequency domain feature makes it less suitable for TSA. Continuous values of the LF/HF ratio must be calculated using a sliding window of inter-beat intervals. This window needs to be of a certain size in order to obtain meaningful results.

For the present study a 30 second window was used. The result is that the continuous LF/HF series is a smooth signal, lacking time resolution. This may not be a problem for studies of psychological variables that change gradually over time, however, musical changes are comparatively frequent. Consequently, no clear relationships between the LF/HF ratio time series and animation cues were observed.

Cross-correlation analyses: Cross correlations were performed as a means of statistically analysing relationships between the musical decision making and cardiac feature time series. This also allowed the investigation of whether the time series exhibited lagged correlations. Initially, the group-averaged time series for percussionists and pianists were used. No consistent correlations between cardiac activity and decision making were observed across all the analyses for averaged time series. Cross-correlations for individuals were then analysed, using only HRV as the cardiac feature. Again, the results did not reveal any consistent trends or correlations either for individuals, or over groups of instrumentalists.

In summary, whilst a number of significant discrete, time-averaged correlations were observed between cardiac and musical decision making measures, the TSA did not reveal any equivalent continuous time series correlations. A potential issue is that there are multiple random and non-random factors that influence both musical decision making and cardiac variables over time. When averaging firstly over two minute periods, and then over groups of participants, random variables will become less influential. However, when looking at continuous values, the influence of these variables is likely to be more prominent. In the case of collaborative musical performance, these variables might be the result of environmental, social, and affective influences; which are difficult to measure and control. Indeed, continuous-time relationships between musical decision making and cardiac activity may only be revealed when such variables can be accounted for.

5.6.3 Experimental Design

With regard to the experimental design, there are clearly a huge number of circumstances that involve co-present musical performance. Each circumstance is defined by an array of variables, such as the instruments played; the musical style; the experience and personalities of the musicians; and the presence or absence of an audience. It is not possible to design an experiment where all of these variables are modified and tested. Consequently, these variables were kept in mind, whilst an attempt was made to design a study that enabled exploratory questions and hypotheses to be tested in a statistically valid manner. Study 1 was highly controlled, in the sense that single drums were used in a laboratory setting, with an experimental task that was not particularly representative of real-world musical interactions. On the other hand, the musicians were free to improvise, entirely of their own accord. In contrast, the present study was designed to be more musically varied and representative or real world musical interactions. At the same time, the decision was made to impose some more restrictions on the improvisation task. There were three main reasons for asking the musicians to create an accompaniment to an animation. The first was to provide some guidance and purpose to the musicians' improvisations; the second was to encourage a greater amount non-verbal communication between the musicians; and the third was to introduce something else for the musicians to focus their attention on, other than their instruments and fellow performers. By using the same animation for all experimental dyads, it was hoped that this would act as a controlling factor, facilitating analyses across dyads. Regarding the first and second objectives, the animation was successful in both attracting the attention of the musicians and encouraging them to work collaboratively. This was evident from observing their performances and their discussions during the warm up performances. Regarding the third objective, observations also indicated that the musicians were attentive to the events and storyline of the animation, and adjusted their playing in response. This was also supported by the analyses of the two time series measures of musical decision making: musical change points (CPs) and information content (IC). Figure 5.13 shows that both measures appear to increase on, or around the animation cue points. This suggests that participants were introducing musical changes that coincided with the animation cues.

Recruiting pairs of musicians who were unknown to each other was also a conscious design decision. The intention was to avoid introducing uncontrollable variables pertaining to the relationships between musicians and their past experiences of playing together. Musicians will often perform composed music in ensembles with people they do not know well. However, with the exception of some particular scenarios, such as open 'jam sessions'[6], they are less likely to engage in improvised performance with people they do not know. Consequently, it was difficult to predict how the musicians would react to being asked to improvise with someone they had never met before. Throughout the experimental sessions the researcher was surprised at how quickly and well the musicians settled into this scenario; and at the quality of the music they created.

[6] A musical activity where people come together to create music by improvising, without substantial preparation.

5.7 Conclusions

The purpose of the second study was to test the LuminUs: investigating the influence of its feedback upon collaborative musical improvisation; whilst also investigating the potential for cardiac activity to be used in the future provision of feedback relating to musical decision making. The study was designed to test a number of specific exploratory questions and hypotheses. To conclude the study, the findings are summarised relative to each of these questions and hypotheses.

On the Effects of the LuminUs (LQ)

LQ1 Does the LuminUs influence the musical outcomes of the collaborations? The findings did not indicate that the LuminUs had a significant influence upon the musical outcomes. The measures of musical outcomes consisted of three items from the self-report questionnaire: i) 'I liked the music we created'; ii) 'I was satisfied with my musical contributions', and iii) the ratings of the creativity of each accompaniment. On average, participants agreed with the first two items more during the conditions where they had some form of LuminUs feedback. However, this was not statistically significant. No positive effects were found of the LuminUs feedback on the creativity scores. This suggests that the LuminUs may have had some positive influence on the participants' perceived musical outcomes. However, this influence was not substantial. Furthermore, there is a likelihood that the 'novelty factor' of the LuminUs may have biased the participants' ratings.

LQ2 Does the LuminUs influence self-reported measures of the musicians' enjoyment and interaction with their co-performer? The findings did not indicate that self-reported measures of enjoyment and interaction were significantly influenced by the LuminUs feedback. However, the statements – 'I enjoyed the session', 'I felt connected to the other musician', 'I felt engaged with the other musician', and 'the other musician and I performed well as a pair' – all received higher average agreement ratings in the two LuminUs feedback conditions. Again, there is a possibility that these results were biased by the novelty of experiencing the technology for the first time.

LQ3 Does the gaze feedback influence the number of glances exchanged between musicians during collaborative interaction? Findings showed that the gaze feedback condition had significant positive effects on the number of glances exchanged between musicians. The proportion of reciprocated glances (within 5 seconds) was also found to be significantly higher when both participants were receiving gaze feedback, and the quantity of feedback was significantly positively correlated with the number of glances.

LQ4 **By facilitating communication between musicians, does the gaze feedback influence the number of musical changes made during a composition?** The findings did not indicate that the gaze feedback had any influence upon the automatically extracted measures of musical change - change point (CP) count and mean information content (IC).

LQ5 **Does the motion feedback have an impact upon the overall amount of motion during collaborative interaction?** No evidence was found to suggest that the motion feedback had a measurable influence upon the musicians' body motion.

LQ6 **Does the motion feedback stimulate increased awareness of the other participant, as indicated through more glancing at the other?** The number of glances exchanged between participants were found to be significantly higher in the conditions where they both had motion feedback. Moderate, but not statistically significant effects were also seen in the conditions where only one participant was receiving motion feedback.

On Cardiac Activity and Musical Decision Making (CH)

CH1 **Levels of the LF/HF ratio of heart rate variability will be correlated with the number of musical decisions:** Some support for this hypothesis was found. Specifically, the mean LF/HF ratio for pianists was significantly correlated with the mean IC of their accompaniments. A positive effect was also found for percussionists, however this was not significant ($t = 1.246$, $p = .220$). For the change point count, no effect upon LF/HF ratio was found for pianists, and a weak positive effect was found for percussionists ($t = 1.421$, $p = .164$). For the continuous time analyses, it was noted that LF/HF ratio has poor time resolution due to the use of windowing to calculate frequency domain values. This may have contributed to the fact that no time series relationships were observed between decision making and the LF/HF ratio.

cH2 Heart rate variability will be negatively correlated with the number of musical changes: The findings provide limited evidence to support this hypothesis. The findings showed moderate negative effects of mean IC and CP count upon mean HRV for pianists. However, these effects were not significant. For the time series analyses, HRV appeared to be decreasing or within local minima at the animation cue points. These points also appeared to coincide with increases in the number of musical changes. However, the cross-correlation analyses did not show any consistent correlations between musical change or IC series and the HRV series.

cH3 Musical decisions will tend to coincide with heart rate extrema: Limited evidence was found to support this hypothesis. The discrete analyses showed that the number of HR extrema were significantly correlated with the mean IC for pianists. However, this does not necessarily imply that the extrema coincided with changes. The time series analyses indicated that extrema for the pianists' mean HR series tended to coincide with the animation cue points. Again, this does not necessarily imply that musical changes coincided with HR extrema. Furthermore, since this result was solely based upon visual analyses, it is only suggestive of a relationship.

5.8 Summary

This study investigated the LuminUs: a novel device for providing real-time interpersonal feedback to collaborating musicians. To the researcher's knowledge, this is the first time that anyone has used eye-tracking to provide interacting musicians with real-time feedback upon the glances of their co-performers. Broad, quantitative analyses were performed, indicating that the LuminUs had a significant influence upon the glancing behaviours of the musicians.

The present study also built upon the findings of Study 1 and investigated specific hypotheses concerning relationships between cardiac activity and musical decision making. In order to do this, new methods were adopted and developed for automatically extracting quantitative measures of musical decision making from MIDI data. This is the first time, to the researcher's knowledge, that anyone has used such measures in the analysis of improvised musical performance. Correlations were analysed between these novel measures and a broad range of cardiac features. Again, this is not something that is believed to have been undertaken in previous studies. Furthermore, a combination of both discrete and continuous time techniques were used to perform the analyses. The description, application, and results of these analyses could be beneficial to those looking to undertake similar studies in the future. The results revealed potentially meaningful relationships between cardiac activity and musical decision making, predominantly from the discrete, time-averaged analyses. However, it is worth stressing that, given the lack of previous research, these results should be viewed with caution. In particular, they highlight the need to take into account other factors that influence cardiac activity during musical performance. The results indicate that some of these factors are dependent on the instrument being played. They may also be dependent on the measurement time period (e.g. fatigue related changes in cardiac activity).

The present study had a number of limitations. Firstly, it only allowed the participants to test the LuminUs within a specific performance scenario, with a co-performer who they had not previously performed with. Secondly, the experiment only provided a short amount of time for the musicians to get used to using the LuminUs. In addition to this, the collected data were predominantly quantitative, and lacked a more qualitative perspective on the musicians' experiences with the LuminUs. These limitations pointed towards the need for further work to investigate the sustained use of the LuminUs in more natural settings.

Chapter 6

Study 3

Investigating the LuminUs in Practice

The previous chapter reported a study in which the LuminUs was put to the test in a controlled environment. The findings of the study indicated that the gaze and motion feedback of the LuminUs had significant effects upon the gaze interactions between musicians. Whilst the controlled design of this study enabled statistically valid, quantitative analyses to be performed, it also placed some limitations upon the findings. In particular, the LuminUs was tested within a restricted performance scenario, where participants were paired with people they had not previously met. This means that the findings may not be generalisable to other situations and contexts. Furthermore, the participants were only given the opportunity to use the LuminUs for a short time period; leading to potential 'novelty factor' influences.

To address these limitations, this chapter reports a qualitative study, which investigated the use of the LuminUs by experienced musical duos over the course of three rehearsals and a performance. The main purpose of this study was to gain insights into musicians' use of the LuminUs in more naturalistic settings, over an extended period of time. It was also hoped that the findings would provide a qualitative perspective on the predominately quantitative findings obtained in both of the preceding studies. The data from the present study consisted of video observations and interview data, and is analysed using thematic analysis. Five themes are established, which are evidenced and discussed in detail.

6.1 Aims

The main aim of the present study was to supplement the findings reported in the previous chapter with a qualitative investigation into musicians' use of the LuminUs over a longer time period. The study was also seen as an opportunity to reflect upon the research throughout this thesis from a qualitative perspective. Specific aims are discussed below, according to four categories: **appropriation**, **meaning**, **impact**, and **design**.

Appropriation: From a technology standpoint, appropriation refers to the way that a particular technology is adopted and adapted by a particular user, or set of users, according to their particular circumstances, interpretations and needs. This aim concerns an interest in the appropriation of the LuminUs by pairs of musicians. The justification for this aim surrounds the fact that Study 2 was carried out in a controlled setting. This had a number of implications. Firstly, the participants were not able to discuss their use of the LuminUs during the experiment. Secondly, they were only using the device for a relatively short period of time ($\sim$ 1 hr.). Thirdly, the experimental task – composing an accompaniment to an animation with a musician they had not previously met – was not necessarily representative of their real-world collaborative music making experiences. These factors enabled the data to be subjected to valid quantitative analyses. However, they also limited the scope for drawing broader conclusions and insights from the study. One particular example of this is the 'novelty factor' issue, highlighted in Section 5.7.

The intended outcomes of investigating appropriation revolved around gaining a better understanding of the musicians' use of the LuminUs in 'real world' circumstances such as rehearsing, composing, and performing. In particular, an intention was to learn about new, and unforeseen ways in which the musicians might use the device in their everyday circumstances. Additionally, it was hoped that important features of collaborative music making interactions could be highlighted by observing how the musicians chose to appropriate the LuminUs.

Meaning: This aim concerns the way in which musicians interpret and draw meaning from the feedback provided by the LuminUs. Study 2 showed that the LuminUs had a significant impact upon gaze interactions between musicians during collaborative music making. However, the data did not provide substantial insight into the causal processes and mechanisms that may have led to these results. For example, it was found that providing gaze feedback to musicians led to increased glancing between them, but it was not possible to provide empirical evidence to indicate why this result was obtained.

By observing and discussing the use of the LuminUs with musicians, it was hoped that specific insights would be gained into how the musicians interpret the feedback provided by the LuminUs. This would also facilitate an interrogation of the meaning behind their responses to the feedback. Additionally, it was hoped that discussions with the musicians would shed light upon subjective interpretations of words, such as 'engagement' and 'connection', which were used to gather self-report data in studies 1 and 2.

Impact: An important aim of this study was to investigate various types of impact that the LuminUs might have on collaborative music making. These include both short term impacts, such as being able to communicate more successfully during a performance; and longer term impacts, such as changing a musician's approach to collaboration. Again, a key motivation for this aim was the 'novelty factor' issue associated with Study 2, whereby the impact of the LuminUs may have been related to it being something that the musicians had not experienced before. Furthermore, Study 2 only collected subjective measures of impact through a post-task questionnaire, with no opportunity for participants to provide more in-depth feedback about their experience of using the LuminUs and its impact upon them. By undertaking a qualitative investigation of the LuminUs it was hoped that specific and detailed accounts of its impact would be evidenced. These data would then serve as a qualitative contrast to the quantitatively established impacts evidenced in Study 2.

Design: One of the underlying motivations of this thesis is to contribute towards the application of new sensor technologies in a collaborative music making context. In this sense, the LuminUs represents an experimental step towards the realisation of real-world devices and applications. Musicians were not given the opportunity to provide feedback on the design of the LuminUs during Study 2. Consequently, an aim of the present study was to obtain detailed findings relating to design considerations and suggestions for improvement to the LuminUs. It was also hoped that these findings would contribute towards an informed discussion and reflection upon the potential roles for new affective/social signal-based technologies in collaborative music making.

6.2 Review of Qualitative Research Methods

In contrast to the preceding studies, where quantitative research methods were used, the present study adopts a qualitative approach. The design of this study was informed by existing qualitative methods, which are briefly reviewed in this section.

Qualitative research has been generally defined as "a naturalistic, interpretative approach concerned with understanding the meanings which people attach to phenomena (actions, decisions, beliefs, values etc.) within their social worlds" (Snape and Spencer, 2003, p. 3). Two opposing paradigms exist concerning the nature and acquisition of knowledge through qualitative research (Snape and Spencer, 2003; Blandford, 2013; Lapan et al., 2011). The first is *positivism*, an objective and law-governed approach, which holds that knowledge exists independently of the researchers seeking to acquire it. The second is *interpretivism*, which posits that knowledge and meaning are 'constructed' through the views and interpretations of participants and researchers. It is worth noting that, despite the contrast between these two stances, it is possible to undertake mixed-paradigm research by combining both positivist and interpretivist approaches (Lapan et al., 2011). This can involve the coupling of statistically-driven, quantitative research, which is inherently positivist, with interpretivist, qualitative research.

In HCI-related fields, qualitative methods are becoming increasingly popular as a means of evaluating new systems and interfaces (Adams et al., 2008). These methods are often drawn from long-established methods that have their roots in the social sciences (Anderson, 1994; Klein and Myers, 1999). One of the challenges for researchers undertaking qualitative studies is deciding which method is best suited to their particular research objectives (Blandford, 2013). In reality this may involve tailoring a combination of existing methods to suit the particular requirements and circumstances of the study. Blandford (2013) uses the term semi-structured qualitative study (SSQS) to describe various qualitative approaches that are commonly used in HCI research. These methods involve the collection and analyses of data in a way that is adaptable and open to interpretation, whilst still maintaining a degree of structure. Furthermore, SSQSs are suitable for investigating a range of research questions. Two important methodological considerations in the planning of any SSQS are data collection, and analysis techniques. The following section reviews three data collection methods that are relevant to the present study: semi-structured interviews, observations, and focus groups. Section 6.2.2 then addresses specific analysis methods.

6.2.1 Data Collection Methods

A variety of methods exist for data collection during SSQSs (Blandford, 2013). When selecting methods for data collection the researcher should also take into account the ways in which they intend to analyse their data. In some cases analysis of the data will take place concurrently with the data collection process. Methods that are specifically relevant to the present study are described below. As previously mentioned, a combination of these methods may be adopted within the design of a single SSQS.

Semi-structured Interviews

Interviews enable researchers to investigate the opinions and subjective experiences of individuals. A semi-structured interview involves interviewing participants using a set of prepared questions or topics as a guide. This means that the interviewer does not necessarily need to adhere to a strict structure, enabling them to probe and explore themes as they arise during the interview. It is important for the researcher to consider their active role in the data collection, and the impact that this could have on the outcomes. For example, researchers should attempt to use non-leading questions and should avoid disclosing personal information, which may bias the responses of the participants (Snape and Spencer, 2003).

Observations

Observational data collection can be separated into *non-participant* and *participant* methods of observation (Flick, 2009). In the former, the researcher takes a passive role and attempts to minimise their influence upon the observations. In the latter, the researcher steps into the observational arena as an active participant, and consequently influences the observations. Participant observation is an important aspect of ethnography - an approach to qualitative research that has its origins in social anthropology (Snape and Spencer, 2003; Ritchie, 2003). Ethnography involves making detailed observations of people in everyday, naturally occurring settings (Randall and Rouncefield, 2013), and has been adopted as an approach to studying and designing HCI (Grudin, 2012; Anderson, 1994).

During observational data collection, the researcher utilises all the senses available to them in order to record any data that may be of relevance to their topic of research. Other considerations are whether the participants should be made aware that they are being observed, and whether the observations should be made in a controlled or natural environment (Flick, 2009). Discreet observation can be achieved using video cameras.

However, research ethics normally require that participants are made aware if they are being videotaped.

Focus Groups

During a focus group the researcher takes a less active role, serving to mediate discussions within a group of participants (Blandford, 2013). Focus groups allow verbal data to be collected through more naturalistic forms of communication between participants. This is useful for investigating context-specific attitudes and practices. Focus groups also enable researchers to validate the views of individuals by observing the extent to which these views are accepted by the wider group. Researchers should take into account that the outcomes of focus group data collection can be greatly influenced by relationships and dynamics within the group.

6.2.2 Analysis Methods

The data collection methods discussed in the previous section often result in text-based data; such as transcribed interviews, written observations, and documented sources. A common initial step in the analysis of such data is to code them (Lapan et al., 2011; Howitt and Cramer, 2011). *Coding* involves the researcher labelling individual segments of the data with words or short sentences that are in some way representative of their content. The types of codes that the researcher assigns to the data will depend upon the aims of their study, their understanding of the subject, and their methodological choices. For example, codes could refer to topics, places, people, or abstract ideas and concepts. Codes are usually developed and refined in an iterative manner, which enables the researcher to establish a more concise representation of their data; facilitating the identification of key ideas and themes, whilst still maintaining a connection to detailed source data. Three commonly adopted methods for the analysis of discourse-based, qualitative data are grounded theory, discourse analysis, and thematic analysis. These methods are described below. In each case, specific coding styles and requirements are defined.

Grounded Theory

Grounded theory is a methodological framework for developing theories from qualitative data. Consequently, it incorporates methods and guidelines, not just for the analysis of qualitative data, but also for the planning and collection of these data. Numerous methods for data collection can be employed, including those mentioned above. Two important features of grounded theory are that it is *inductive* and *iterative* (Lapan et al., 2011). Inductive refers to the fact that knowledge is constructed 'from the ground up'; meaning that researchers develop theories from the data, rather than commencing the research with preconceived hypotheses (Charmaz, 2006). Grounded theory is iterative in the sense that analysis is performed concurrently to data gathering, with the analytical outcomes feeding back into the ongoing data collection. During grounded theory analysis the researcher works from initial coding of the data towards more focused coding and the generation of categories. Throughout this process constant comparisons are made between codes, data, and emerging categories, in order to lead towards the development of theories. A key method within grounded theory is theoretical sampling. This involves selecting where and what data should be collected, based upon the theories that emerge during the study. When the collection and analysis of new data ceases to lead to new theoretical insights then an end point is reached, referred to as 'theoretical saturation' (Charmaz, 2006).

Discourse Analysis

Discourse analysis encompasses a range of approaches to understanding discourse (i.e. speech or text) as social action. A central concept is that language is used as a tool for doing things, such as constructing and interpreting meaning (Howitt and Cramer, 2011). The source data for discourse analysis is normally written text transcribed from interviews or conversations. However, there are a wide range of procedures for undertaking discourse analysis (Howitt and Cramer, 2011). Stowell et al. (2009) adopted a structured discourse analysis method for the evaluation of a live music making interface. This involved the itemisation of transcribed conversations into specific objects, such as 'synthesiser' or 'participant'. Roles and descriptions were then ascribed to these objects, allowing the researchers to reconstruct their relationships and identify underlying properties of the interactions. It should be noted that discourse analysis has limitations in that it is difficult to learn and there is a lack of consensus surrounding appropriate methods.

Thematic Analysis

Thematic analysis centres around the identification of themes in data via a process of coding, similar to that used in grounded theory. However, unlike grounded theory, it does not require iterative data collection and theoretical sampling, meaning that it can be applied to pre-existing data sets. Furthermore, thematic analysis is not reliant upon specialised theoretical frameworks, making it more widely accessible to qualitative researchers than methods like discourse analysis. Consequently, the flexibility afforded by thematic analysis has contributed towards it being one of the most commonly adopted methods for qualitative analysis (Braun and Clarke, 2006; Howitt and Cramer, 2011). This flexibility does not limit the power of thematic analysis as a tool that can provide detailed and complex accounts of qualitative data (Braun and Clarke, 2006).

Data for thematic analysis may be gathered through a combination of methods, such as those discussed in Section 6.2.1. Having familiarised themselves with the data, the researcher then goes through a process of coding, as introduced at the start of this section. Coding in thematic analysis can be 'data-driven' or 'theory-driven' (Braun and Clarke, 2006; Howitt and Cramer, 2011). In the former, the coding is primarily dictated by the content of the data. In the latter, the influence for the coding derives from the theoretical standpoint of the researcher and their specific research questions.

Once all the data have been coded, the researcher searches for themes, into which the codes can be categorised. This is normally an iterative process of selecting candidate themes and checking to see how well they represent both the codes and the original data. Eventually, a set of themes are decided upon and are clearly defined with the support of extracts from the data.

6.3 Research Design and Data Collection

A qualitative study was designed based upon the aims set out at the beginning of this chapter. These aims centre upon the use of the LuminUs by musicians in more naturalistic settings. Consequently, a longitudinal study design was chosen, whereby established musical duos were asked to attend four separate sessions with the LuminUs, each lasting roughly 90 minutes. Data were collected from video observations, semi-structured interviews, feedback surveys, and a final group discussion. This section describes the study design and data collection in detail.

6.3.1 Software and Hardware

During Study 2 participants had no control over which type of feedback the LuminUs was providing: gaze or motion. Consequently, modifications had to be made to the LuminUs software to enable participants to freely switch between the two types of feedback during the present study. A basic software interface was designed (see Fig. 6.1) that enabled each participant to select which type of feedback they wanted to transmit to their co-performer's LuminUs. The main part of this interface consisted of

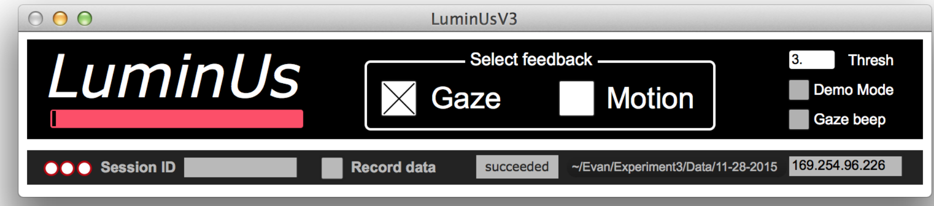


Figure 6.1: The LuminUs software interface used in Study 3.

two check boxes for selecting the feedback type. Additional check boxes allowed the participants and researcher to put the LuminUs into a demo mode, and to trigger an audio tone (beep) whenever the participant gazed at the eye-tracking marker. These functions were only used for testing that the devices were set up correctly. A numerical input box, labelled 'Thresh', provided the option to modify the threshold distance for the gaze detection; however this did not end up being adjusted during the study. The lower rectangular pane on the software interface was designed for use by the researcher, allowing them to verify that the software was receiving data. It also provided the option to record data, and to configure the IP address used to transmit control signals to the LuminUs over a wireless network.

The LuminUs software was further modified to facilitate the use of two distinct gaze detection markers for each participant. In Study 2 a single type of marker, consisting of concentric black circles (see Fig. 4.12 on page 106), was attached to the eye-tracking headsets. This meant that in the gaze feedback conditions a participant's LuminUs would light up when their co-performer was looking at them. In the present study markers were still attached to the headsets; however, the participants were also provided with an extra marker to place wherever they chose. The reason for this was to provide the participants with added scope for appropriating the LuminUs. It was also thought that observing where participants chose to place the second marker would reveal important visual reference points in the musicians' environment. In order for the eye-tracking software to distinguish between two separate markers, the type of marker had to be changed. Square fiducial markers, such as those shown in Fig. 6.2, were used, since the Pupil software already had the capability to recognise them. It was also necessary to modify the light feedback provided by the LuminUs so that the musicians could identify which marker their co-performer was looking at. This was achieved by programming the lights so that they turned increasingly red when gaze towards

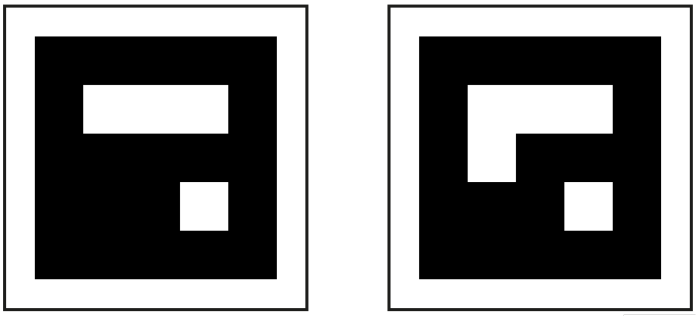


Figure 6.2: Examples of two of the fiducial markers used for gaze detection.

the headset marker was detected; and increasingly blue when gaze towards the extra marker was detected.

The extra markers for each participant were printed onto squares of card, 10 cm × 10 cm in size. These were then provided with a choice of holders: either a weighted stand, or a flexible clamp arm. The LuminUs devices were mounted on flexible arms, which were then attached to microphone boom stands. These features meant that participants had the freedom to situate the markers and LuminUs devices in a variety of positions. Figure 6.3 shows an example of one of the set ups used in the study, with the LuminUs to the left of the drum kit and the extra marker clamped in the centre.

Figure 6.3: An image showing how one of the participants set up their equipment during the study, with the LuminUs positioned on a flexible stand to the left, and the extra marker clamped to the centre of the kit.

6.3.2 Participants

Four experienced musical duos were recruited to take part in this study. Unlike Study 2, where participants in each dyad had not previously met, this study required that each pair of musicians had prior experience of playing together as a duo. This decision was made due to a desire to investigate the ways in which musical dyads might integrate the LuminUs into their existing arrangements, and the influences that it might have on their established collaborative interactions. It was also deemed more representative of real-world music making scenarios, where musicians would tend to know each other.

Participants who had taken part in Study 2 were invited to take part in this follow up study. One of the four groups consisted of a duo where both of the members had participated in Study 2. This was due to the fact that they had attended the same London music school, from which participants were originally recruited. Two of the groups contained one musician who had participated in Study 2 and one who had not. The fourth duo were recruited from a general advertisement and neither of them had participated in Study 2. The fact that some participants already had experience with the LuminUs was not seen as an issue, given that this study was not targeted towards controlled, statistical analyses. Furthermore, the selection of these participants was entirely based upon their availability to participate during the specified study dates, and was not in any way influenced by their previous data or responses to the LuminUs during Study 2. Each participant was offered £80 for their time and travel expenses, on the condition that they attended all four sessions. Details of the four dyads are provided in Table 6.1. As can be seen from this table, the musicians covered a range of musical styles and had a combined experience of 132 years. Ethical approval was acquired for the study (QMREC1547a).

Table 6.1: Group and participant details for Study 3.

Group	**Style**	**Par.**	**Age**	**Sex**	**Instrument**	**Exp.**	**S2?**
1	Jazz	1	25	M	Drums	8	Y
		2	25	M	Piano	18	Y
2	Folk/blues /funk	1	22	M	Bass guitar	17	N
		2	22	M	Guitar	12	Y
3	Rock/pop /classical	1	20	F	Drums	7	N
		2	39	F	Piano	30	Y
4	Rock	1	34	M	Guitar	19	N
		2	32	M	Drums/drum machine	21	N
Abbreviations: Par. (participant), Exp. (experience, in years), S2? (participant in study 2? Yes (Y) or no (N))							

6.3.3 Sessions

Each group attended four sessions, which were held in the performance and rehearsal spaces at Queen Mary University of London. The sessions were designed so that three aspects of musical collaboration could be investigated: rehearsal, composition, and performance. The first session required the musicians to rehearse a single piece of music of their choice. The second a third sessions required them to compose and rehearse a new piece, at least three minutes in length. The final session was a group session where each group performed their composed piece in front of the other study participants. Each session was roughly 90 minutes long, comprising an hour for rehearsing/composing/performance, followed by roughly 20 minutes for a semi-structured interview. Ten minutes were factored in for additional activities, such as setting up instruments.

Upon arrival at the first session, the participants were given a written and verbal description of the study and were asked to sign consent forms. They were then shown how to calibrate the gaze tracking headset and how to use the software interface to switch between the two types of LuminUs feedback. Following this, participants were given a demonstration of both the gaze and motion feedback functions of the LuminUs. This was a purely functional demonstration. No instructions or suggestions were provided as to *how* the participants were expected to use the LuminUs during the sessions. Participants were aware that the purpose of the experiment was to investigate their use of the LuminUs. However, they were merely told that the device was there for them to use if they chose to. Throughout this introductory period participants were free to ask any questions. Once the participants and researcher were satisfied with the set up, the participants were instructed to spend the remainder of the hour rehearsing a single piece of music. The researcher then left the room, informing the participants that he would be sat outside if they had any further questions or problems during the session. Once the hour was up, the researcher entered the room to inform the participants. The session then concluded with a semi-structured interview (see Section 6.3.4), conducted in the rehearsal space with both of the participants at the same time.

The second and third sessions followed a similar format to the first, except that the groups were instructed to spend their time composing and rehearsing a short (3-5 min) piece of music for the final performance session. The final session was held in the performance space. Chairs and tables were set out for the participants to watch the performances, and a stage was set up at one end of the room, with stage lighting and a PA system. Each group then took turns in performing their composed piece. Following each performance, groups were instructed to provide a brief verbal description, summarising their use of the LuminUs. Each participant was also given a brief feedback form (see Section 6.3.4) to complete after their performance. This session concluded with a group discussion, lasting roughly 20 minutes.

6.3.4 Data Collection

Two primary methods of data collection were adopted for this study: non-participant observation, and semi-structured interviews. Non-participant observation was chosen in order to avoid influencing the participants during the sessions. This was achieved using a single, discreetly positioned video camera. Semi-structured interviews were chosen to enable the collection of data relating to specific aspects of the musicians' use of the LuminUs. The interviews were held at the end of each session and were recorded on both a video camera and a stand-alone audio recorder. Nazroo and Arthur (2003) stress the importance of planning and preparing for interviews, even when a flexible interview format is used. They suggest that the general structure of an interview should be split into three stages. The first stage should consist of introductory and easy opening questions. This gives the researcher an opportunity to gather some contextual and background information, and it allows the participant to be eased into the interview. The second stage leads into more in-depth questions and forms the core part of the interview. Nazroo and Arthur suggest that questions about experiences, behaviours, and circumstances should precede those about motivations and attitudes. The final stage involves winding down the interview, during which it can be a good idea to get the participant to summarise their views or experiences. These stages provide a broad structure to the interview, but Nazroo and Arthur also recommend the use of a topic guide for planning and conducting interviews. The topic guide acts as a flexible agenda for the interview, providing a list of topics to be covered, which may also include specific questions and follow-up questions. From a practical standpoint, topic guides work best when they consist of brief words or phrases that serve as prompts to the interviewer. In the case of this study, topic guides were developed for each session. An example of the topic guide for the first session is provided in Table 6.2. After the first session, topic guides were partially customised for each group in order to address or follow up on specific observations or topics from the previous session. The researcher was aware that the participants' knowledge of his involvement in the development of the LuminUs might bias their responses in favour of the device. Therefore, participants were explicitly encouraged to provide critical feedback during the interviews.

The group interview at the end of the final performance session was conducted

Table 6.2: An example of one of the interview topic guides used in this study, including the rough time (in minutes) allocated to each set of questions and follow up questions (> and >>).

Time	Question	>	>>
	Introduction:		
	How often play?		
2	How long known?	How meet?	
	Gig often?	What venues?	
	General LuminUs questions:		
2	Tech used before?	Describe?	
	Problems?	Overcome/how resolved?	
	Practical questions:		
	Feedback used most?	Why?	
		Instrument/role?	
	Marker placement?	Why?	
5		How gaze used?	
	Accelerometer placement?	Motion important?	What does it communicate?
	Where LuminUs?	Why?	Meaning of orientation?
		Instrument/role?	
	Functional questions:		
	Discussions had?	Why?	Agree/disagreements?
5	Partner reaction to feedback?	How recognised?	How expected?
	Actual reaction?	Why?	
	Evaluation questions:		
	Influence on playing?	Why/how?	
5	What enjoy performing together?	Did the LuminUs influence?	How?
	How establish/maintain engagement?	Did the LuminUs influence?	How?
	Why is it important to engage?	Points when more/less important?	
1	***Winding down:***		
	Changes for next session?		

in the style of a focus group (see Section 6.2.1), where the researcher attempted to mediate discussion between the musicians. A short follow up survey was also given to participants during this session (see Appendix C.4). The survey questions were as follows:

- Name:
- Age:
- How many years have you been playing?
- What style of music did you play this evening?
- Did you end up using the technology in your rehearsals? (YES / NO)
  - If YES - how did you use it? If NO - why didn't you use it?
- Did you use the technology differently during the performance? YES / NO
  - If YES - how did it differ?
- What, if anything, did you like about the technology?
- What, if anything, did you dislike about the technology?

6.4 Thematic Analysis

As highlighted in Section 6.2.2, thematic analysis offers researchers a flexible method for qualitative analysis, which still has the potential to provide a thorough and complex account of the data (Braun and Clarke, 2006). This section provides a detailed report of the thematic analysis of interview, survey, and observational data collected during this study.

6.4.1 Method

Braun and Clarke (2006) describe six phases of thematic analysis, which were used to guide the analysis reported in this section. The phases are as follows:

Becoming familiar with the data This is a vital first stage in the analysis process, which involves the researcher familiarising themselves with the data. One of the first ways of achieving this is for the researcher to be actively involved in the task of transcribing data. Once the data are converted to text then the entire data set should be read and re-read. Notes should also be taken at this stage, marking out early ideas.

Initial coding This stage involves systematically going through the entire data set and assigning initial codes to segments of the data (see Section 6.2.2 for more details on coding). As previously mentioned, coding may be data-driven, or theory-driven.

Searching for themes Once the data have been coded, the researcher attempts to group and categorise these codes into broader themes. At this stage, visual representations, referred to as 'thematic maps', may be used to help gain an overview of relationships and patterns in the coded data. Some themes may come directly from codes, and it is also possible to create a hierarchy of sub-themes.

Reviewing themes The researcher reviews and refines the candidate themes obtained in the previous phase. This is a two level process. Firstly, the coded extracts within each individual theme are reviewed to ensure that they form a coherent pattern. If they do not, it may be necessary to modify the name of the candidate theme, or to re-assign coded extracts to an alternative, or new theme. Secondly, the entire data set should be re-read to assess whether the thematic map provides a suitable representation of the data set as a whole. During this process the researcher is also advised to re-code any segments of data that were missed during the initial coding. Again, this may also result in the identification of new themes, and the rejection or modification of others. These steps can be repeated until the researcher is satisfied with their thematic map.

Defining and naming themes This stage requires the researcher to formulate a clear and detailed definition of each theme. This involves reviewing the extracts within each theme and structuring a coherent account; identifying interesting aspects of the theme in relation to the overall aims and research questions posed by the study. A written analysis of each theme should be provided. The use of sub-themes can be helpful for breaking down particularly large and complex themes.

Producing a report The report should provide an interesting and detailed account of the story told by the data. Evidence should be presented for each theme, using a sufficient number of data extracts to convince the reader of the validity of the analysis. Alongside this, the researcher should provide their own narrative, relating their analysis to the research questions and existing literature.

6.4.2 Data

The majority of the data for thematic analysis derived from the video and audio recordings of the study sessions and interviews. These were manually transcribed by the researcher using NVivo[1]. The transcriptions consisted of verbatim accounts of all the dialogue in the recordings, except for that which was entirely unrelated to the study. Non-speech utterances, such as verbalisations of musical phrases, were approximated where appropriate. Non-verbal communicative acts, such as gestures, were only transcribed when they were required to make sense of the dialogue. For example:

PARTICIPANT 2: If I want to make something go louder I would start to play louder, but also [*mimes playing guitar gradually louder, does head and facial expressions*]

The raw interview data comprised 13 audio recordings (3 interviews per group $\times$ 4 groups + 1 group interview), lasting a total of three hours, fifty seven minutes ($M = 18$ min). The transcribed interview data contained a total of 1458 lines of transcribed dialogue ($M = 112$) with a total word count of 36715 words ($M = 2824$).

The raw session data comprised 12 videos[2], lasting at total of ten hours and fifty-four minutes ($M = 54$ min). The transcribed session data contained a total of 508 lines of dialogue ($M = 42$) and a total word count of 6040 words ($M = 503$). Eight handwritten survey responses were also transcribed.

6.4.3 Coding

Coding of data for thematic analysis can be performed at two different 'levels': a semantic level, and a latent level (Braun and Clarke, 2006). In this study, coding of the transcripts was performed on a semantic level. This means that the codes were derived from explicit meanings within the data, rather than latent, underlying ideas and concepts. As previously discussed, coding can also be 'data-driven' or 'theory-driven' (see Section 6.2.2). A data-driven approach was used in this study, such that the selection of codes was not constrained by ideas and preconceptions concerning the use of the LuminUs.

Following familiarisation with the data, initial coding was performed. An example of a segment of coded transcript is shown in Table 6.3. The entire data-set was then re-read so that codes that had emerged late in the first coding iteration could be applied, where necessary, to earlier segments of the transcript data. This process generated 150 distinct codes.

[1] http://www.qsrinternational.com/

[2] The third session for group 4 was not recorded due to equipment failure.

Table 6.3: An example of the coding of an interview transcript.

Speaker	**Transcript**	**Codes**
Interviewer	So, could each of you just summarise for me how you've ended up using it [the LuminUs]?	
Participant 2	Probably the same use I did since the first moment, just knowing when he looks at me, so we can communicate better.	Communication, gaze, usage, awareness of other, use over time.
Participant 1	Yeah.	
Participant 2	I think it's as simple as that.	Simple use.
Participant 2	And cues.	Cueing.
Participant 1	Pretty much, it's just getting cues...	Cueing.
Participant 2	Again, if we were playing something that was completely arranged, maybe I wouldn't use it at all. Maybe I wouldn't need it. I would focus a lot on my thing, while paying attention, listening to him, but not really care to look at him. Because I can listen to him, so I can focus my eyes here. But now that we have to look at each other or the tune cannot be played, I'm actually using this.	No use for tech, gaze, NVC, situation, musical style, hearing the other, usage.

6.4.4 Theme Development

Based upon a preliminary categorization of codes, the following candidate themes were selected: adopting the technology; context; expression and communication; imagining the technology; musical affordances; roles and responsibilities; and working as a group. Coding and theme details were exported from NVivo and imported into Gephi[3], an open-source graph visualisation platform (see Appendix C.3 for a full list of the source data). This facilitated the creation of a thematic map by semi-automatically grouping codes around candidate themes, and scaling the code names according to their

[3]https://gephi.org/

prevalence in the data. The resulting thematic map is shown in Fig. 6.4. ‘Adopting the technology’ and ‘expression and communication’ were clearly the most prevalent themes, in terms of the number of codes they referenced.

The seven candidate themes were then reviewed and refined. This involved scrutinising how well the coded extracts matched their themes, and how well the themes matched the data set as a whole. During this process, some codes were merged and others were split in order to improve the coherence of the themes. For example, the ‘confidence’ code was originally assigned to the ‘adopting the technology’ theme, since it coded extracts where participants had expressed lack of confidence in the LuminUs. However, upon inspection it was found that this code had also been used to code extracts that related to the underlying confidence of the musician. Consequently, the code was split into two new codes – ‘confidence of musician’ and ‘trust and confidence’ – which were assigned to the ‘working as a group’ and ‘adopting the technology’ themes respectively. New codes were also created in order to group together codes that had similar meanings, and codes that were so specific that they only referenced single extracts in the data. For example, the codes ‘looking to pass the ball’ and ‘looking for appraisal’ were both grouped under a new ‘functions of gaze’ code. As the process of refinement progressed, some of these new codes became sub-themes.

The themes themselves were also redefined and, where possible, re-organised using sub-themes. Most notably, the ‘adopting the technology’ theme was merged with the ‘imagining the technology’ theme and renamed ‘technology’. Due to the size of this new theme, five ‘technology’ sub-themes were created: design, impact, potential, problems, and usage. Despite its comparable size, the ‘expression and communication’ theme could not be segmented into sub-themes. This was predominantly due to the multiplicity of functions, styles, and meanings of expressive and communicative acts. Other notable revisions to the themes were the renaming of ‘musical affordances’ to ‘aspects of music’, and the merging of ‘roles and responsibilities’ and ‘working as a group’ to form the new theme, ‘group attributes’.

The process of reviewing and refining the codes and themes continued until the existing themes provided a comprehensive and coherent account of the data, and no additional themes were emerging. At this point, a final thematic map was developed, which is shown in Fig. 6.5. The final themes are: aspects of music, context, expression and communication, group attributes, and technology. Each of these themes is described in more detail in the following section.

Figure 6.4: The candidate thematic map. Candidate themes are shown in green and codes are shown in blue.

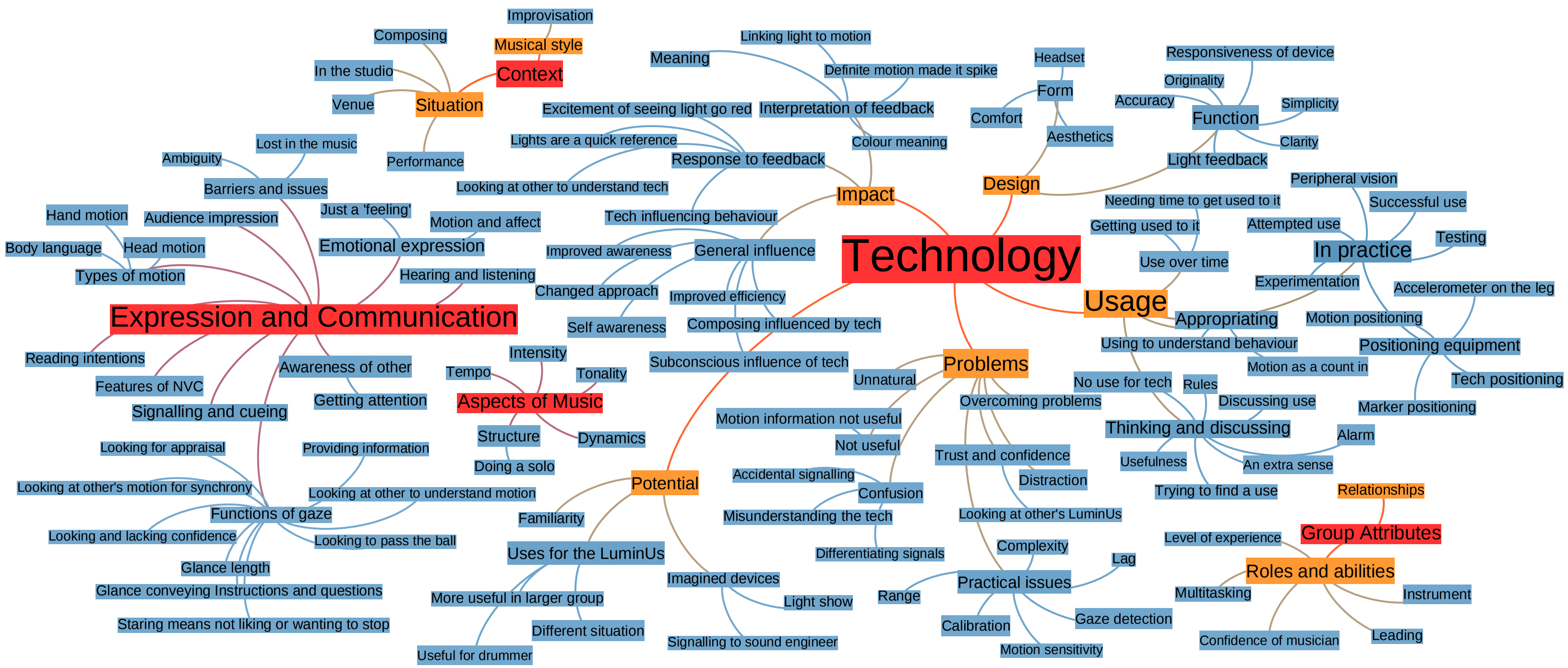


Figure 6.5: The final thematic map. Themes are shown in red, sub-themes in amber, and codes in blue. Lines indicate which codes and sub-themes contributed toward the main themes.

6.5 Final Themes

In the previous section five themes were identified: expression and communication; group attributes; technology; aspects of music; and context. In this section each of the themes are described and evidenced in detail, using observations and excerpts[4] from the interview transcripts.

6.5.1 Expression and Communication

The theme of expression and communication was one of the most prominent themes to emerge from the analyses. Clearly this relates closely to the fact that the LuminUs is designed to aid and enhance non-verbal communication (NVC), and this was a prevalent topic throughout the discussions with musicians. It is worth noting that there is a distinct reason why 'expression and communication' was selected as a title for this theme, rather than 'non-verbal communication'. Communication implies the transfer of information between two people. Expression can be seen as a key element within the process of communication, whereby thoughts or feelings are expressed as visible actions. However, if an expression is not observed then no communication occurs. Therefore, whilst communication implies expression, the opposite is not necessarily true. As will be seen in examples throughout this section, musicians discuss how they use expression in a functional, communicative way, such as for cueing different segments of a piece. They also discuss more affective aspects of expression, such as the way it reflects their feelings and engagement with the music. In this latter case, the expression is not necessarily intended as a conscious communicative act. However, through an awareness of their co-performers, musicians describe how they attend to, and interpret these expressions.

As previously discussed, it was not possible to find a clear way of breaking this large theme into distinct sub-themes. However, for the sake of clarity, this section is split into two sub-sections. The first concerns the various **functions and styles** of expression and communication that were discussed and displayed by the musicians: *why* and *what* do people express and communicate? Following this is a discussion of the **modes and methods** of expression and communication: *how* do people express and communicate? Some of the transcript excerpts have relevance within both of these sub-sections, so cross-referencing is used where appropriate.

[4] A note on the formatting of excerpts: text inside square brackets is used for clarification, and to replace participant names; italic text inside square brackets is used to describe actions in cases where they assist in the interpretation of the excerpt.

Expression and Communication: Functions and Styles

The dominant form of expression and communication that emerged from the analyses was 'signalling and cueing'. Signalling and cueing can be seen as *functional* acts of expression, where the musician has an explicit message that they wish to communicate. 'Cueing', in a musical sense, specifically refers to the act of signalling a specific action, such as changing section or ending a piece. In this sense, cueing tends to revolve around actions that occur at discrete points in time. In the following excerpt a participant describes various uses for cueing:

PARTICIPANT 2: For example, going into a solo, usually the band leader would look at the person who's going to do the solo. It's like passing the ball. Let's say I'm the leader, I would play my solo and then when I'm done I would look at the guy so that he takes the ball and then he runs with his solo. When you are changing keys, or when you're ending, or when you're going into a funny section..., yeah, so everybody looks.
(Excerpt 1: Group 3, Interview 3)

For other references to cueing, see excerpts 26 on page 208; 24 on page 207; 77 on page 225; 45 on page 215; 62 on page 219; and 66 on page 220.

Musicians also used and discussed forms of signalling that were more continuous in time, such as varying the volume of musical segments. The prevalence of signalling and cueing themes were influenced by the ways in which musicians saw the LuminUs as a device that could assist them in these acts. This is discussed later in this chapter (see Section 6.5.3).

Expression and communication of *emotion* was also a prevalent theme. In comparison to signalling and cueing, emotional communication does not tend to convey explicit intentions, or result in tangible outcomes; making it harder to define and observe. Indeed, when talking about emotional expression and communication, the musicians often used vague definitions such as "it's just a feeling" or "the energy in the room":

PARTICIPANT 2: There are times when we were playing now when we both knew that we both liked it. [...] Because you feel like an energy in the room when you're playing something that you really like. (Excerpt 2: Group 4, Interview 1)

PARTICIPANT 1: You can feel the mood of the music, [...] It's hard to explain.
PARTICIPANT 2: Obviously sometimes you make really silly decisions and the other one, [...] maybe he doesn't show it straight away, [...] but...
PARTICIPANT 1: You feel that [the music] has suddenly gone down in energy.
INTERVIEWER: Is that just a thing that you hear?
PARTICIPANT 2: Yeah.

PARTICIPANT 1: You hear and you feel, yeah. (Excerpt 3: Group 1, Interview 3)

A potential factor in the use of vague descriptions may be the fact that emotional communication clearly comprises many modes and forms of expression, as exhibited by the following excerpt:

INTERVIEWER: What are the factors that have made you feel "this isn't working"?
PARTICIPANT 1: I think you feel it, [. . .] the energy, the body posture, the facial expressions. And you hear it in the playing. [. . .]
PARTICIPANT 2: And if they're not enjoying it, I guess you can feel it.
(Excerpt 4: Group 2, Interview 3)

During performances, expressions and communication between musicians can also have an influence upon audience perceptions. Musicians discussed how they are conscious of the ways in which their expressions can impact an audience:

PARTICIPANT 2: As a performer, you can't be lost in the music. It's not good. Because you always have to maintain this thing with the audience, you see. And if you're lost, you've lost them. (Excerpt 5: Group 3, Interview 3)

PARTICIPANT 2: You care a lot about stage presence [. . .] and do things on stage to keep things interesting. When someone is head banging a lot, you want to head bang as well, [. . .]. It's mainly for the show. (Excerpt 6: Group 1, Interview 3)

PARTICIPANT 2: I think you can tell if somebody's into the music or not. If they're not they wouldn't move as much. But then again, there's some people that stand really still, but you see their facial expression.
PARTICIPANT 1: But then again, that could be part of the performance as well, so it wouldn't necessarily be obvious.
PARTICIPANT 2: But then we are marked on that kind of thing, aren't we? You have to be more expressive. (Excerpt 7: Group 3, Interview 1)

Participants also discussed how, as audience members, the communication between musicians makes an impression upon them. For example, in the following excerpt a participant describes how it can be more impressive to witness musicians doing something in synchrony without any visible communication occurring between them:

GROUP 1, PARTICIPANT 2: The point is not to see a performance and understand why. [. . .] It's meant to be like, they do this rest, [. . .] and you don't understand how they did it so good without even looking at each other. (Excerpt 8: Group Discussion)

However, as the following excerpt indicates, this is not always the case, and may be somewhat dependent upon the style of music:

GROUP 1, PARTICIPANT 1: I think it also depends on the style of music. For example, in jazz [...] I would prefer that the members that are playing are looking and there's interaction [...] and you can see how the music is developing through the interaction. But then there's other stuff, which is very well prepared and which is, like, it's "trigger, bam, bam, da, da, da" [*gestures sudden events with hand*]. And that would be much more better if they are not like [*mimes signalling with gaze and nodding*].
(Excerpt 9: Group Discussion)

Evidently, expression and communication are integral to functional, affective, and aesthetic aspects of collaborative music making. Furthermore, despite addressing these aspects separately, it is clear that they are not mutually exclusive; the act of communicating cues between musicians could also have an impact upon affective and aesthetic aspects of audience perception.

Expression and Communication: Modes

Musicians referred to a wide range of modes of expression and communication during collaborative music making. Due to the fact that the LuminUs is designed to provide feedback on gaze and motion, these were the most frequently discussed modes.

With respect to **gaze**, the analyses led to the identification of prevalent functions and features of gaze in collaborative music making. The first notable feature of gaze is that it is used as both a means of expressing information, and receiving information. With regard to expressing information, gazes were rarely discussed or enacted without the involvement of additional modes of expression, such as movements of the head, eyebrows, and mouth:

INTERVIEWER: So, you say you'd normally use a glance to indicate these kind of things, like when someone comes in...

PARTICIPANT 2: Yeah, or a nod. Or you would use some kind of body motion, or your eyes. (Excerpt 10: Group 3, Interview 1)

INTERVIEWER: If you're improvising and you decide you're going to make a change [...] how would glancing factor in to you changing? Would you use glances during that process?

PARTICIPANT 1: I guess so. I think it would be a mix between what you play – like, you signal it in your playing, maybe – and also, like, get the attention of other people.

PARTICIPANT 2: Like, if I want to make something that should go louder I would start to play louder, but also, like [*mimes playing guitar gradually louder, does head and facial expressions*].

PARTICIPANT 1: [. . .] so it's like a mix between your ears and visual things.
(Excerpt 11: Group 2, Interview 2)

INTERVIEWER: Say you're playing with a pianist and he or she changes what they're playing, would you..., to change your drumming accordingly, would you be looking towards them, or just listening to that change?

PARTICIPANT 1: A mixture. [. . .] I would be looking out for a signal from them if they're saying "we're doing this now" [*does head and eye expression*], I'd be, like [*does thumbs up, raises head and opens eyes wider*]. Or if they were literally just going crazy, I'd just be listening and just going along with it. (Excerpt 12: Group 3, Interview 3)

As a mode of expression, gaze alone clearly has very few degrees of freedom. The length, and quantity of glances were two features that were discussed in relation to communication mediated through gaze:

INTERVIEWER: So, if you just glanced at someone briefly, compared to staring at them, do you think there's a difference in what you're trying to communicate that relates to the length of time that you would look across at someone?

PARTICIPANT 1: Well, if someone's looking at you like..., [*does long stare*] you're obviously going to be like "oh, they're actually looking at me for a while!"

PARTICIPANT 2: You don't generally look at somebody like that, unless they've done something really bad. Usually when you're playing together it's just a very quick glance and the person knows what you want to do, especially if you play together quite often.
(Excerpt 13: Group 3, Interview 1)

PARTICIPANT 2: I'm trying to think of a time when the feedback made me stop. I guess probably when I did notice [Participant 1 was] looking at me quite a lot, I was like, "okay, maybe now it's time to stop". (Excerpt 14: Group 4, Interview 1)

Regarding the use of gaze for receiving information such as signals, cues and emotional expressions (discussed earlier in this section), visual attention is clearly important. In addition to this, musicians also described looking at their co-performers to gain information about timing, and for feedback on musical contributions:

PARTICIPANT 1: Yeah, like, if you're on one side of the percussion pit, or whatever, and there's a timpani bang that's got to be exactly in time with the cymbal crash at the other side, you'd have to be like [*mimes looking over and doing pronounced arm motion*], so you get it together. (Excerpt 15: Group 3, Interview 3)

PARTICIPANT 2: When you're composing and you have to arrange and make a form, then you always need to look and be sure about how many times everything is going to happen. And also to get the "okay, go" from the other player, the "is it OK, what I'm doing; do you like it?" (Excerpt 16: Group 1, Interview 3)

Regarding **motion**, various types of motion were discussed pertaining to different parts of the body. Of these, head motion was most frequently discussed. A possible explanation for this relates to the fact that head motion does not tend to be restricted by the requirements of playing an instrument. As pointed out by a musician in the following excerpt, this makes the head "mutual to mostly all instruments":

PARTICIPANT 1: I think most motion when you're playing, [...] with regards to intensity and stuff like that, would be mostly in the head...

PARTICIPANT 2: In the head, yeah. I mean, if it goes really intense you will head bang. If it's really calm, you'd be completely still. [...]

INTERVIEWER: Does it depend on the instrument?

PARTICIPANT 1: Actually, I think the head is the one part of the body which is mutual to mostly all instruments. If you know what I mean?

PARTICIPANT 2: Yeah, yeah. Except the singers. (Excerpt 17: Group 1, Interview 1)

In practice, 'nodding' was referred to as a particular case of instructive head motion:

PARTICIPANT 1: When would you look at me if you wanted to end? [...] When would I give you the nod? (Excerpt 18: Group 3, Session 1)

Most discussion of motion did not refer to specific gestural motions, and instead referred to more generalised descriptions of motion, such as the amount, or frequency of movement. This motion was linked to the musicians' feelings towards the music being played (see also Excerpt 7 on page 203):

PARTICIPANT 1: Well, also you were saying you'd be interested in the motion to see whether I was responding to the music. (Excerpt 19: Group 4, Interview 1)

Expressive motion can be directly linked to the motions of playing one's instrument. This has relevance to the next mode of communication: **hearing**. Musicians frequently discussed how expressive and communicative attributes of the musical collaboration were attended to through listening to the music itself:

INTERVIEWER: If it's feeling good [...] how do you get a sense of that? How do you get a sense of the other person's experience of the music?

PARTICIPANT 2: Well, you kind of see it in their face...

PARTICIPANT 1: In their smile. You can hear it as well, in their playing, I think. At least, since I know your playing, [...] like, when we do function gigs and we're depressed, I can hear it in his playing. [...]

INTERVIEWER: Can you put a finger on it?

PARTICIPANT 1: I don't know. You're actually, like, giving yourself into...

PARTICIPANT 2: Into the music.
PARTICIPANT 1: ...instead of just, like, being a robot.
(Excerpt 20: Group 2, Interview 1)

PARTICIPANT 2: Still, force is something you immediately understand by hearing the other one, but movement you can't understand until you see him.
(Excerpt 21: Group 1, Session 1)

As can be seen from some of the statements above, musicians draw links between what they can hear and certain features of the performer's motion. Other modes of expression and communication that musicians discussed in less detail included posture (or body language), and facial expressions.

Musicians also spoke of barriers and issues concerning expression and communication. These included physical barriers and relationship-related issues, which are discussed in sections 6.5.5 and 6.5.2 respectively. A foundation of communication is awareness. Having an awareness of others in an ongoing aspect of musical collaboration, but musicians frequently discussed the need to get the attention of a co-performer:

PARTICIPANT 1: ...sometimes you're in your own world when you play, and then you just need the attention of the other guy. (Excerpt 22: Group 2, Interview 2)

PARTICIPANT 1: Because some people are lost in the music, and you want to get their attention, [...] you're thinking "I really want to communicate with the drummer to tell them this..." (Excerpt 23: Group 3, Interview 3)

PARTICIPANT 2: Sometimes you try to look at everyone and cue them, and someone might not be looking at you, and you get miscommunication.
(Excerpt 24: Group 1, Interview 1)

Another issue is that expressions often carry a degree of ambiguity, meaning that they might not always be interpreted correctly, as highlighted in the following excerpt:

PARTICIPANT 2: One of the ideas we had was maybe we can use [the LuminUs] to signal certain things in the music [...] Because sometimes it's a bit confusing when you do this [*makes expression with face*], does it mean "go on" or does it mean "end"?
(Excerpt 25: Group 3, Interview 1)

Summary: The theme of expression and communication incorporates many important aspects of the interactions between collaborating musicians, as well as their individual behaviours and experiences. It can be direct and purposeful, as in the case of signalling and cueing. In the case of emotional expression and communication, it is ever-present, often occurring in subtle forms and through both conscious, and subconscious processes. In performance situations expression and communication can be influenced by the presence of an audience. In each case, musicians utilise the expressive and communicative modalities available to them, such as glances, motions, gestures and hearing. As modes of communication, these are not infallible. They can be influenced by human factors, such as misinterpretation; and by external factors, such as physical barriers. Indeed, each of the themes to emerge from the analyses contains attributes that influence expression and communication. These are highlighted, where possible, in the following sections.

6.5.2 Group Attributes

Group attributes are, in most cases, attributes that exist prior to musical collaboration commencing. For example, the instruments played by the musicians, and their relationships with the other musicians. The group attributes theme is separated into two sub themes: **roles and abilities**, and **relationships**.

Group Attributes: Roles and Abilities

The roles and abilities sub-theme refers to the specific attributes and responsibilities of individuals within a group. These may be interrelated. For example, a musician will have certain requirements and restrictions relating to the instrument that they play, such as a drummer needing to be sat down; and they will also have specific responsibilities, such as dictating the tempo. These factors contribute to the ways in which a particular musician will interact with other members of the group, as indicated by the following excerpt:

INTERVIEWER: In a band situation, who would you be looking at the most?

PARTICIPANT 1: Well, it depends [...] we do function gigs and then it's very important with the singer, to make sure that we all go to the same sections at the same time. [...] But maybe the drummer as well [...]

PARTICIPANT 2: I think for you it's probably drummer isn't it? As a bass player. [...] When we play function stuff we look at each other... for the next part.

(Excerpt 26: Group 2, Interview 2)

Some attributes of a musician's role may vary *during* collaborative music making. One such attribute is leadership. A previous excerpt has already indicated how leadership responsibilities may be passed between members of a group during playing (see Excerpt 1 on page 202). During this study, the roles of the musicians also influenced and placed restrictions upon the ways in which they used the technology:

PARTICIPANT 2: I don't know if the other people that do this have a bit more multitasking skills, but I play the drums, I've already got a lot to do, and to then really handle a lot of complex feedback, I don't think would have worked.

(Excerpt 27: Group 4, Interview 1)

PARTICIPANT 2: I can't find a way to use the motion, to be honest. Especially with the instrument I play. I don't know. I'm sitting down.

(Excerpt 28: Group 1, Interview 3)

The abilities of individuals within a group also have a bearing upon collaborative music making. The level of experience of the musicians may dictate their ability to recognise and act upon signals and cues; and their ability to play their instrument whilst concentrating on other things can influence their awareness of co-performers:

PARTICIPANT 2: What I do, I need communication with the singers.
INTERVIEWER: Do they normally lead that line of communication?
PARTICIPANT 2: If they're not very good musicians then yes, [...] sometimes if they're not very confident they'll look at me like this [*mimics looking for help*], which is bad.

(Excerpt 29: Group 3, Interview 3)

Group Attributes: Relationships

The relationships between musicians have a definite impact on their collaboration. Existing experience of playing music together, and familiarity with a musician's behaviours can have an influence on expression and communication, as implied in the following excerpt:

PARTICIPANT 2: It would be good to try [the motion feedback] next time. Then again, I'll probably move loads and [Participant 1] just tend[s] to be... [*laughs*]. I'm in a band with someone who's very static and stares down all the time.

(Excerpt 30: Group 4, Interview 1)

PARTICIPANT 1: If you're used to playing with certain people..., used to the way you communicate with that person..., like, you might arrange your own signals and things.

(Excerpt 31: Group 3, Interview 1)

The relationship between musicians does not necessarily depend upon them having had previous contact. In such circumstances, aspects of the relationship become apparent in the moment of collaboration, and may be harder to define, as discussed in the following excerpt:

PARTICIPANT 1: I think there are certain people who you can improvise with, and there are certain people who you can't improvise with. [...]

PARTICIPANT 2: But this doesn't mean that to improvise with someone you have to know him. It can be a random musician that you meet for the first time, and suddenly you lock. (Excerpt 32: Group 1, Interview 1)

Summary: Group attributes incorporate a variety of factors that determine the ways in which a particular group of musicians will interact and collaborate together to make music. Some of these attributes are a function of the individual musicians, such as individual abilities and instruments played. However, once these individuals become part of a group, their individual attributes can no longer be considered independently of the group, and must be considered in relation to the attributes of the other group members.

6.5.3 Technology

The technology theme is a substantial one, which covers all aspects relating to the use and discussion of the LuminUs and other technologies. The large quantity of data relating to this theme led to it being separated into five sub themes: **usage**, **impact**, **problems**, **design**, and **potential**. Each of these sub-themes is discussed separately, although there is clearly some crossover between them, which is addressed in the discussion section of this chapter.

Technology: Usage

As it suggests, the usage sub-theme refers to discussions and actions of the musicians as they went about adopting and using the LuminUs during the sessions.

Most groups commenced their first session with the LuminUs by getting straight into playing music. This was usually interrupted by an initial period of discussing and thinking about how they would use the device, as well as testing its capabilities. The following excerpt is an example of such a discussion:

PARTICIPANT 2: I really like the idea of knowing that if I'm playing, that you're looking at me. Because I normally can't see that. And then you've got that, kind of, alarm. Whereas, I've noticed it does get to your peripheral vision quite well, doesn't it? [...]

PARTICIPANT 1: So, I need to put my marker...

PARTICIPANT 2: You need to put your marker somewhere I can look at it easily to let you know. So probably just in front of you. [...] It should go blue when that thing... So if you put it a little bit closer, like, pop it on the floor down there, or something, or...

PARTICIPANT 1: [*moves marker towards Participant 2*] Yeah, now it will get it. [...] So if I keep moving back. [*walks backwards with the marker*] You're still looking at it, yeah?

PARTICIPANT 2: Yeah, but no. Maybe you rest it up on [the table].

PARTICIPANT 1: So it only works, I guess, to there.

PARTICIPANT 2: Yeah, so maybe pop it on the top of that door or something? Yeah, so that's just my alarm to you to. [...] Hopefully, when you're playing you should see that in your vision. (Excerpt 33: Group 4, Session 1)

This excerpt also illustrates how the musicians would test the capabilities of the device. In this case they were testing the range of the gaze detection. The amount of testing and experimentation undertaken with the LuminUs varied greatly from group-to-group. Testing of the gaze feedback mainly concerned testing the range of the gaze detection and whether the device was working. Less testing was performed with the motion feedback. Evidently, some musicians felt pressured to find a use for the LuminUs in their playing:

INTERVIEWER: So did you have many discussions? It seems like you talked quite a lot at the start about things, was that related to how you were going to use it?

PARTICIPANT 2: Yeah, we talked a lot, because [Participant 1] said that me *must* find some way to use this and incorporate it into our playing. We were also testing how it works. (Excerpt 34: Group 3, Interview 1)

In practice, two out of the four groups ended up settling upon specific uses for the LuminUs. Group 1 used the gaze tracking as a means of notifying each other when they were being looked at; as a way of improving their cueing abilities. In Group 2 one of the musicians used the motion feedback to indicate how intense he thought the other musician's solo should be, by attaching the accelerometer to his ankle. In this case the motion feedback was visible to both musicians. The second musician in this group used the eye-tracking to cue certain aspects of the music, such as key changes. He would do this by looking at one of the gaze markers so that the LuminUs would light up, giving a signal to the other musician. An unexpected use of the gaze feedback, which was attempted by two of the groups, was using it as a tool for identifying how often certain objects were being looked at. In one case this object was a drum machine; in the other case it was a lyrics sheet:

PARTICIPANT 1: So maybe we should put my eye thing with [the lyrics sheet], because then it can sense when I need to read the lyrics.

PARTICIPANT 2: OK, which is all the time.
PARTICIPANT 1: [*Places the marker down by the lyrics*]

(Excerpt 35: Group 2, Session 1)

A practical aspect of the musicians use of the LuminUs concerned the positioning of the LuminUs devices, as well as the accelerometers and the eye-tracking markers. Participants tended to simply position the LuminUs devices where they were easily visible. One of the drummers moved his device to either side of his drum kit, depending upon which part of the kit he was using most for that particular song, as described in the following excerpt:

PARTICIPANT 1: I moved [the LuminUs] there [by the floor tom]. [. . .] for the only reason that when we were mainly improvising I was using the toms mostly, so I was looking that way, and I couldn't really see it from here [*points to right side of the drum kit*], [. . .] It depends, if it was another song I would probably leave it here.

(Excerpt 36: Group 1, Interview 3)

There did not seem to be any particular consideration for the orientation of the devices. The group who settled upon using the gaze feedback placed their markers on their instruments to supplement those being worn on their eye-tracking headsets. For the other musicians the placement of the markers seemed to be more influenced by the presence of a surface or object somewhere in front of the musician, where the marker could be placed or attached (see Excerpt 33 on the preceding page).

A final feature of the technology usage was the way it developed over the course of the four sessions. Participants discussed a need to "get used to" aspects of the device, such as the feel of the eye-tracking headset and the presence of the feedback. The following excerpts are from discussions after the second and third sessions with the group who settled upon using the gaze feedback:

PARTICIPANT 2: I used it more today than the other time. Because probably you need some time to get used to having this option. To look there and..., kind of, with the peripheral vision. (Excerpt 37: Group 1, Interview 2)

INTERVIEWER: What would you say are the main things you've learnt about the device, or adapted, or changed?
PARTICIPANT 2: I'm paying more attention to it.
PARTICIPANT 1: Yes, actually, I am more as well. I think [. . .] you start not to think about it any more. Like, from the first session we were like,"ah, is it working?" [. . .] But now it's just, like, you start off and you just look there and it's lighting up.

(Excerpt 38: Group 1, Interview 3)

Technology: Impact

The impact sub-theme accounts for the ways in which musicians interpreted and responded to the feedback from the LuminUs, as well as the general influence of using the device during the study sessions. Regarding the gaze feedback, the basic interpretation appeared to be fairly clear to the musicians: when the light comes on it means that the other musician is looking at one of the markers. There was occasionally some confusion over which colour of light responded to which marker (the headset or the moveable marker). In contrast, the motion feedback was far more open to interpretation. In the following excerpt a drummer describes how he interpreted the motion feedback and associated it with different actions, such as using a guitar pedal:

INTERVIEWER: So there wasn't enough variation in the light, compared to the movement?

PARTICIPANT 2: Well, not really, and I think it was kind of distorted. The most input was when you were actually just doing something that was kind of utilitarian [. . .] you had to move a pedal, so you had to walk, and that was a stronger signal than when we were actually playing through something you seemed to like.

(Excerpt 39: Group 4, Interview 2)

A member of a different group described how he interpreted the motion feedback as being too representative of the frequency of movement, and that it should instead be more about the *force* of the pianist's playing:

PARTICIPANT 1: I have a suggestion actually, [. . .] the motion sensor shouldn't be about the frequency of movement, because if you're playing something soft, very fast, [the LuminUs] shows totally red. It should be [. . .] force which you're playing.

(Excerpt 40: Group 1, Session 1)

The vague interpretability of the motion feedback meant that some musicians initially responded to the motion feedback light on the LuminUs by glancing across at their co-performer to see what kind of motions they were doing. This allowed them to establish their own interpretations of what the feedback meant:

INTERVIEWER: Do you think that [the motion feedback light] had any influence on how much you looked at each other?

PARTICIPANT 2: Yeah, I think it did, [. . .] there were definitely times when the light came on and I looked up, but I wasn't sure if I was looking up to see if [Participant 1] was into it, or looking up to see how the device was working. [. . .] That's how I found that point about [Participant 1] walking forwards when he's found an idea he likes. [. . .]

INTERVIEWER: So in terms of looking across [...] would you have done that when it was all lit up, or when it wasn't lit up, or just any point?
PARTICIPANT 2: When it was all lit up. (Excerpt 41: Group 4, Interview 2)

Generally, the response to the feedback was influenced by the ways in which the musicians were attempting to use the LuminUs. For the group using the gaze feedback, the response to the light was simply to look towards the other musician:

INTERVIEWER: How do you respond to the light?
PARTICIPANT 2: Just straight I'm looking at him. Whatever I'm doing I'm looking at him. (Excerpt 42: Group 1, Interview 3)

In the following case the musician described an added pressure associated with the glance when using the gaze feedback:

INTERVIEWER: Did you notice when you looked at that marker, for instance, and [Participant 2] was looking down; did you notice a reaction?
PARTICIPANT 1: It's difficult to say, because I do, kind of, set up to look and focus on [Participant 2], so [...] it's difficult to..., know. I mean, I suppose. Did you feel like you were being looked at [*to Participant 2*]?
PARTICIPANT 2: Yeah, I felt a pressure, yeah definitely. It was definitely more of a pressured glance... (Excerpt 43: Group 4, Interview 1)

For other groups there did not appear to be any response at all:

INTERVIEWER: So when you, say, looked at your marker, did you notice any reaction in [Participant 2]? Or when you looked at her?
PARTICIPANT 1: I don't think so.
INTERVIEWER: And you [*to Participant 2*]?
PARTICIPANT 2: No..., no. (Excerpt 44: Group 3, Interview 1)

In light of the variety of uses, responses and interpretations to the LuminUs feedback, it is not surprising that its general influence upon collaborative music making was also found to vary between the groups. The group who settled upon using the gaze feedback reported that, during the composing stage of the study, they felt the increased awareness provided by the feedback had an influence upon their efficiency:

INTERVIEWER: How did [the LuminUs] factor into the way that you composed?

PARTICIPANT 2: I mean, because we're in front of the instruments when we have to play whatever we compose, maybe it helps on the playing side. Like, we can straight away, kind of, focus on our instruments but at the same time know what's happening. [...] Maybe it made it a little bit faster, the composing. Because of the cues, like we were saying, it's more efficient. You know when the other one is looking at you.

(Excerpt 45: Group 1, Interview 2)

Another group discussed how the study had contributed to their awareness of their motions whilst playing. However, as they highlight in the following excerpt, it was unclear to what extent this influence came from the interview discussions as opposed to the actual use of the LuminUs:

PARTICIPANT 1: This whole process I think is..., [...] we never really spoke about it, and this made us very aware of it, and, sort of, very conscious. And I don't know whether that's necessarily because we spoke about it, or because we had these, sort of, instruments in front of us.

INTERVIEWER: When you say "aware of it", you mean aware of...

PARTICIPANT 1: Well, because we spoke about, you know, that's why we can give each other a signal about whether we're enjoying it, or whether it's locked in. So it's difficult to know whether I felt like I was moving more because we had the [LuminUs], or because we had spoken about it.

INTERVIEWER: Yeah. But you just generally have more awareness of your body language when you're performing? Or..., is that what you're implying?

PARTICIPANT 1: Yeah definitely. (Excerpt 46: Group 4, Interview 3)

None of the groups indicated that they felt the LuminUs had an influence upon the quality of music they were playing and composing. A couple of the groups indicated that their composing was influenced by a desire to find a use for the LuminUs:

SURVEY: *What, if anything, did you like about the technology?*

PARTICIPANT 2: It was exciting, made us work differently in the sense that we tried to make music that could have any use of it. (Excerpt 47: Group 2, Feedback Survey)

PARTICIPANT 1: We should definitely incorporate the technology. So we could, like, signal song changes or something. (Excerpt 48: Group 3, Session 2)

Technology: Problems

This sub-theme covers specific problems that the musicians reported, or encountered in relation to the LuminUs. From a practical standpoint, one of the main issues was the reliability of the eye-tracking. In order for the tracking to work well, the calibration procedure must be performed whilst keeping one's head still. If the eye-tracking headset is subsequently repositioned on the person's head then the accuracy can be lost, requiring calibration to be performed again. Furthermore, the gaze feedback is dependent upon the successful detection of the eye-tracking markers. If a marker is positioned too far away, or if there is a large amount of movement of the marker relative to the headset, then tracking can fail. One musician highlighted how these factors led to him consciously limiting his head movement:

PARTICIPANT 2: With [the eye-tracking headset], the slightest movement, I know that it might affect, or it will affect [the LuminUs]. So that, kind of, holds me back, first in movement, and then secondly just knowing that "oh shit, now I have to go back and re-calibrate". (Excerpt 49: Group 1, Interview 3)

Some musicians established that the range of the marker detection was an issue during their testing of the device, and overcame this by positioning the marker closer to themselves (see Excerpt 33 on page 211). However, the general lack of reliability of the gaze detection led to issues of trust and confidence in the device. Group 3 had particular issues with the gaze feedback not functioning correctly during their first session. This led them to believe that there was a lag in the gaze detection, and to consequently question the usefulness of the device:

PARTICIPANT 1: So, it's not very accurate, because I'm not looking at you.
PARTICIPANT 2: Try mine. I'm going to look away and then suddenly look at you. Look at your thing. [*They test the gaze feedback*]
PARTICIPANT 1: [...] So there's, like, a two second lag. [...]
PARTICIPANT 1: But if there's a lag then it's not really useful, right?
(Excerpt 50: Group 3, Session 1)

For another group, the lack of confidence influenced how they envisaged they could use the device, given more time to experiment with it:

INTERVIEWER: Are there any more ways that you would like to experiment with it, that you haven't had a chance to?
PARTICIPANT 2: I mean, if I knew exactly how the threshold thing works and if I was one hundred percent sure that whenever he looks exactly here, [the LuminUs] goes suddenly [lit up], [...] So I would take the cues from the light, not from him. [...] Now, I can't trust this completely. (Excerpt 51: Group 1, Interview 3)

A related issue was that, when using the gaze feedback as a means of signalling, one of the groups expressed a desire to have some form of confirmation that the device had worked and that the signal had been seen by their co-performer:

PARTICIPANT 2: So, we're not going to use any of the devices. [...] Because, when we tried to do that thing of getting each other's attention with it, I realised that actually what I need is confirmation back that [Participant 1], sort of, had the signal. And without that I didn't feel confident leaving it on. (Excerpt 52: Group 4, Performance)

In comparison, the motion feedback did not have any reliability issues. However, its lack of responsiveness to immediate motion was discussed as a potential issue:

INTERVIEWER: So in that sense do you think [the motion feedback] was conveying some information to you that was useful at all, or not?

PARTICIPANT 2: I don't think so. [...] If it were just a tiny bit more reactive, like a kind of sweet spot between..., I understand that you don't want it to be immediate, because of the lag, but for it to be just that little bit more responsive.

(Excerpt 53: Group 4, Interview 2)

Furthermore, a number of participants expressed that the motion feedback was not particularly useful to them.

PARTICIPANT 2: For me, the motion detector is not as useful as the eye detector, [...] There's not much I can tell her with [the motion detector].

(Excerpt 54: Group 3, Interview 2)

PARTICIPANT 2: I can't find a way to use the motion, to be honest. Especially with the instrument I play (drums). (Excerpt 55: Group 1, Interview 3)

With respect to general problems identified during the study, a few of the musicians discussed that the technology felt unnatural, both in terms of comfort and how they were generally used to interacting:

PARTICIPANT 2: I think the glasses are weird to wear. I wouldn't feel comfortable, in the studio, at home, or...

PARTICIPANT 1: Yeah, like, on stage. [...]

INTERVIEWER: From a comfort point of view, or how you see the thing as well?

PARTICIPANT 2: Yeah, I think comfort. I mean, I wouldn't play with it on stage either, but it didn't feel..., like, it felt like I was some kind of computer thing.

(Excerpt 56: Group 2, Interview 3)

PARTICIPANT 2: It's more instinctive for me to look at [Participant 1] than for me to look at [the marker], you know. It's almost like I have to hold myself back and do what I don't normally do..., to use the device. (Excerpt 57: Group 3, Interview 2)

One group also indicated that they found the technology distracting in a detrimental way. Although this was partly attributed to the device not working correctly, and needing time to get used to it:

PARTICIPANT 2: [The LuminUs] does take away from playing the music, because I'm too busy looking at whether the thing is working or not.

PARTICIPANT 1: I think, if you were used to it, it would obviously be different, because you wouldn't be trying to constantly..., like, see, "is it working, is it not?" But because this is the first time we've done it, obviously it's going to be quite distracting.

(Excerpt 58: Group 3, Interview 1)

SURVEY: *What, if anything, did you dislike about the technology?*

PARTICIPANT 1: Could take away attention of someone's real face, e.g. looking at the light more than the person. (Excerpt 59: Group 3, Feedback Survey)

Technology: Design

Throughout the discussions with musicians, numerous suggestions were made as to how the design of the LuminUs might be modified. The design sub-theme brings together these suggestions. Broadly speaking, they can be further separated according to whether they relate to the **form**, or the **function** of the device.

When questioned about the **form** of the LuminUs, none of the groups made any suggestions or indications that they would modify it in any way. One musician suggested that he would want the LuminUs to be mounted on a clamp stand, so that it could be positioned somewhere more inconspicuous:

INTERVIEWER: What would you change about it? What do you dislike about it? What are the negatives?

PARTICIPANT 1: Does it have to be on a stand like that?

INTERVIEWER: Not particularly, no.

PARTICIPANT 1: So you can, for example, just clamp it like that [*mimics clamping it to the drum kit*]. It would be good there, for example. Just so that it's more inconspicuous as well. (Excerpt 60: Group 1, Interview 3)

As previously discussed, some musicians mentioned that they found the eye-tracking headset uncomfortable (see Excerpt 56 on the preceding page). However, others said that they found it comfortable. In the following excerpt one member of the group discussed how he found the headset (glasses) painful after a while, and suggests something more like a Bluetooth headset. The other member of the group indicates that the headset was fine, but that this was because he is used to wearing glasses:

PARTICIPANT 1: Does it have to be glasses? [...] Can it be something like the Bluetooth thing, which just clips onto your ear? [...]

INTERVIEWER: Potentially, yeah.

PARTICIPANT 1: I'm not used to wearing glasses, and after a while it starts to be a bit [painful]. [...]

PARTICIPANT 2: I have no problem, [...] I wear glasses every day, so for me it's nothing different. (Excerpt 61: Group 1, Interview 3)

None of the musicians made any indication that the accelerometer was uncomfortable to wear.

With regard to **functional** aspects of the LuminUs design, most discussions revolved around two specific design issues concerning the feedback provided by the lights. The first issue is whether the gaze feedback should simply involve a single light going on/off, as opposed to a gradual increase/decrease in light. The second issue concerns the sensitivity of the motion feedback. The following exchange highlights both of these issues and includes a suggestion from one of the musicians as to how it could be possible to use a foot-switch to change between the two types of feedback:

PARTICIPANT 2: I would like the way that [the LuminUs light] goes gradually up. Because if, for example, by mistake he stares for half a second and [the light appears] I'd be like "oh, what happened?" Now I know that [...] it takes one and half seconds to go up, then I know that he's actually watching me.

PARTICIPANT 1: Because, when you look, you always look, like, at the end, when the section is ending, not bang on beat. So, going up gradually gives some time to say "ah, OK". [...]

PARTICIPANT 2: I can think of a useful way, in that discussion, to use the motion sensor. Because, for example, [...] I'm seeing some of his movements and if it's a really big movement I know that this is a cue, even though he's not looking at me. So if the threshold is really small on the motion sensors, so it doesn't catch the small movements of the drums, but it catches the big ones, the whole hand, then this would be useful. So I don't have to look at him and wait for the hand to do the big movement, I just have the sensor here. I'm playing, I can see it. Or I could have a switch, gaze and motion, for my [foot]. So whenever I want the gaze, I put the gaze. Whenever I want the motion, I put the motion. He wears both sensors.

(Excerpt 62: Group 1, Interview 3)

Regarding the gaze feedback issue, one musician suggested that by simply having the light go on and off it would make the LuminUs less distracting:

PARTICIPANT 2: I didn't really see the point of [the light] decreasing and increasing. I know for motion it could show the intensity of motion [...] but for gaze, she's either looking at me or she's not, and I find that cool down period, or the build up period, not useful and also distracting, because then you're waiting for it to come up.
(Excerpt 63: Group 3, Interview 1)

However, another musician simply felt that the decreasing of the lights should occur more quickly:

PARTICIPANT 1: I think it would be good if – when someone's not looking at you [*signals lights changing quickly*] – rather than going down slowly. If you know what I mean.[...] Because then it's like, "is he still looking?" (Excerpt 64: Group 1, Session 3)

There was also the suggestion that, in situations involving more than two musicians, the colour of the gaze feedback could indicate which musician was looking at you:

PARTICIPANT 1: I think it would be very useful like that, if each band member had their own colour. [...] So it's like "okay, someone's looking at me", and if it's red it's John, and if it blue it's that guy. (Excerpt 65: Group 1, Interview 2)

The issue of whether to have gradual lighting or a simple on/off display is also dependent upon the type of signal or cue that the musicians might be communicating with the device, as discussed by the musician in the following excerpt:

INTERVIEWER: Do you think it's necessary to have this, kind of, gradual increase?
PARTICIPANT 1: Well, it depends on how you would use it. As a cue thing then maybe you wouldn't need the gradual thing, because then you'd just want, like, a light – "ah, okay, it's the next section". But if it's a thing with intensity, or volume, or speed, or something like that - then of course, because then you can see it building [...] and then you can take it higher and higher and play more and more intense.
(Excerpt 66: Group 2, Interview 2)

Regarding the sensitivity issues for the design of the motion feedback, we have already seen how one suggestion simply involved using a threshold so that only large motions would trigger the feedback (see Excerpt 62 on the previous page). Another issue was related to the fact that the motion data are averaged over time, and therefore not particularly sensitive to sudden movement. Modifications were made to the LuminUs software in order to allow the members of Group 4 to adjust the sensitivity and responsiveness of the motion feedback during their final rehearsal session. This appeared to have a positive impact on their use of the device, as discussed below:

PARTICIPANT 2: By changing the responsiveness to it, I think that helped quite a bit, because it was..., it was more responsive and [...] there seemed to be less feedback from you moving around, like, back and forth, today. [...] So, it sort of uncluttered it and meant that I could actually see what was you actually being into the music, rather than just moving back and forth. (Excerpt 67: Group 4, Interview 3)

Another suggestion was more related to the sensitivity of the motion feedback to different types of motion:

GROUP 1, PARTICIPANT 1: I think it would be good to have the differentiation of colours with the motion sensor. Because at the moment, [...] it just defines the quickness of motion, it doesn't define short motions or long motions. You know what I mean? So if it would be like, for example, turns red when there's very short motions and then it's like, a calmer blue, or something like that when it's like, longer motions. That would help identify. (Excerpt 68: Group Discussion)

Technology: Potential

Throughout the study musicians discussed situations and circumstances where they felt that the LuminUs could have potential use. A common suggestion was that the device would be more useful with larger groups of musicians:

PARTICIPANT 1: So, in that way, it could be good if we were a big band. Because then sometimes it can be difficult, like, to get everyone's attention.
(Excerpt 69: Group 2, Interview 1)

PARTICIPANT 1: I think it could be useful. But I think this particular setting, where we're in a small band and everyone's close, is, kind of, out of context I think. But it could be useful in orchestral settings and big shows, or bigger gigs where it's like..., communication is difficult because there's a big space [...] and there's lots of people.
(Excerpt 70: Group 3, Interview 3)

Musicians also referred to specific *situations* where they thought the device would be more useful (see also Excerpt 76 on page 224):

INTERVIEWER: So if you were playing with someone completely new, what kind of things would you want to see from them that would give you a sense of their engagement?

PARTICIPANT 1: I think it would be good, like, with the motion thing, [...] if it could monitor intensity, because that way you can, in a way, feel where he's going with it, and you can adjust what you're doing. I mean, to be fair, if you have a lot of experience, you will know where he's going, you can hear it. But I think that would be helpful.

PARTICIPANT 2: I mean, a vital part of improvisation is looking at each other, so maybe it wouldn't be a bad idea if, I don't know, in the future you try and put it in two rooms – the musicians in different rooms – and have only the machine to communicate. [...]

INTERVIEWER: Do you think that has relevance [...] in a real life situation?

PARTICIPANT 1: Well, actually, in the studio it could work [...] Because in the studio, for example, you have a drum booth, just so that you don't have bleeding into other microphones. You usually have musicians playing at different times. But if the studio has separate rooms and you have [the LuminUs] it could...

INTERVIEWER: Yeah, so you're talking about the gaze?

PARTICIPANT 2: Gaze and motion, yeah. [...] Everything that can give you information. Because if you take away our eye contact it's a big minus.

(Excerpt 71: Group 1, Interview 1)

Another suggestion related to specifically to the potential usefulness for drummers:

PARTICIPANT 1: I think maybe this would be the most useful tool for a drummer, because you can't move around, you can't, like, [...] you know, it's like you're behind this wall of sound that you make. So I think for a drummer it would be..., because then you could get the attention of everybody, with just a light.

(Excerpt 72: Group 2, Interview 1)

One of the more imaginative suggestions was that the device could be used to trigger stage lighting, allowing the musicians to control the light show with their eyes:

PARTICIPANT 1: I was thinking we should use the lights for, like, stage lighting [*laughter*]. No, but seriously. Like, have that as, like a big, like, stage light. And when you want to change the mood of the song [...] you could be like, [*looks suddenly towards marker*] something intense has got to happen, [...] and when I look over here it's pink, when I look over here it's blue. (Excerpt 73: Group 3, Session 1)

Summary: Technological aspects of musicians' experiences with the LuminUs have been discussed under five sub-themes: usage, impact, problems, design, and potential. During use, musicians experimented with the device and tested its capabilities. Two out of the four participant groups settled on specific uses for the technology. The impact of the LuminUs also varied between groups; with musicians reporting multiple interpretations and responses to the gaze and motion feedback. In particular, some musicians indicated that the device had a positive impact upon their efficiency and awareness; whilst others felt pressured to find a use for it.

Problems with the device predominately related to the sensitivity and unreliability of the gaze feedback. This led to issues concerning lack of trust and confidence in the device. For the motion feedback, responsiveness was highlighted as an issue. Some musicians also indicated that the technology felt unnatural and was uncomfortable to wear. With regard to the design sub-theme, topics mainly concerned the *functioning* of the light feedback, rather than the *form* of the LuminUs. Finally, potential uses for the device were highlighted, predominately in relation to specific circumstances and situations.

6.5.4 Aspects of Music

The aspects of music theme refers to the fundamental properties of the music that musicians have control over. Five aspects emerged from the analyses: dynamics, intensity, structure, tempo, and tonality. In many cases these aspects were discussed in relation to the use of the LuminUs, discussed in Section 6.5.3. Therefore, references to excerpts in that section are provided in order to support the observations.

Dynamic properties of the music constitute the loudness, or volume over time. 'Intensity' is a word that was used frequently in the discussions with musicians (see excerpts 17 on page 206; 66 on page 220; 71 on the previous page; and 73 on the preceding page). It is closely related to volume (see excerpts 40 on page 213 and 66 on page 220). However, musicians implied something distinct about intensity, relating not just to the volume, but also to the speed and content of the music, as highlighted by the following excerpt:

PARTICIPANT 2: It's different again in every instrument. For me, for example, intense would be playing loads of notes together and with a weird rhythm and..., But in every instrument it can be different. Maybe in guitar, for example, it would be like doing crazy chords far away from each other, would come out as really intense.

(Excerpt 74: Group 1, Interview 1)

Structure refers to the way in which musical phrases, or segments, are organised in time. For example, a basic song might consist of a sequence of verses with a chorus placed between each of them. A common structural element that musicians discussed was the solo - a section of music where a single musician is either playing alone, or is the centre of attention (see excerpts 1 on page 202; and 77 on page 225). Tempo and tonality are used in the traditional sense, referring to the speed and pitch of the music respectively (see excerpts 66 on page 220; 40 on page 213; and 1 on page 202).

Summary: Whilst this theme may seem unsubstantial in comparison to the earlier themes, it is nevertheless important. The aspects of music have a direct bearing upon the styles of expression and non-verbal communication that musicians use in order to effect change in the music they are making. For example, instantaneous structural changes are signalled by short-term actions, such as glancing; whilst gradual volume changes are communicated through motion. This also influences the ways in which musicians use the LuminUs.

6.5.5 Context

In contrast to the aspects of music theme, which concerns low-level musical properties, the context theme relates to **situational** and **stylistic** aspects of musical collaboration.

The **situational** component of this theme takes into account the location and circumstances surrounding the musical collaboration. These include physical attributes relating to the venue and the positioning of the musicians on a stage. They also include factors relating to the type of event, such as a performance or recording session. Musicians referred to numerous situations when talking about their experiences of collaborative music making, and their thoughts surrounding the potential uses for the LuminUs. Specific situations that multiple musicians referred to included performances, being in a recording studio, and being in an orchestral pit. In the following excerpts participants discuss situations in which the positioning of the musicians has a negative influence on communication between them:

PARTICIPANT 2: I've done a gig recently, [...] the stage had two drum kits and two stands of keys. [...] I had my back to the drummer and the drums were looking the other way, so it was a really strange set up. [...] I could not look at the drums at all. [...] So I thought, like, if only I could have [the LuminUs]. It was the first time that I actually thought "wow, this would be really useful here".

(Excerpt 75: Group 1, Interview 1)

PARTICIPANT 1: We were thinking it would be useful in situations where you can't see each other. Like, if you're in a room, or facing one direction.

INTERVIEWER: Are those situations common?

PARTICIPANT 1: Yeah, they could be. Often being in a pit or something, where you're crammed into a tiny space. Or in a studio, where you're all separated in individual little booths. Actually, that's a common one, in the recording studio when you've got one person that side, in a room over there, and your drums in a separate room, but you're recording at the same time, it might provide a way to communicate.

(Excerpt 76: Group 3, Interview 1)

The second component of the musical context theme is the **style of music** (or genre) being played. Excerpt 9 on page 204 highlights how the musical style influenced the type of interaction that a musician would wish to observe as an audience member. In relation to the LuminUs, musicians also discussed various ways in which the musical style influences aspects of the collaboration:

PARTICIPANT 1: [The LuminUs] could work really well in a blues band, but then it would be in a different way than, like, a..., I don't know, classical, pop kind of thing.

INTERVIEWER: Could you give me an example of how that would be different, maybe, between those two styles?

PARTICIPANT 1: Because, I mean, blues is very cue based. Like, it's really, like, you have the solo and then you build up. (Excerpt 77: Group 2, Interview 1)

Summary: Again, this theme is less substantial than earlier themes in this chapter. However, context plays an important role in determining the nature and limitations of musical interactions; which subsequently has an influence upon forms of non-verbal communication, and the potential for technological interventions. For example, musicians playing in a darkened room may be hindered in their ability to use visual forms of communication, which could make a device like the LuminUs more useful.

6.6 Discussion

In contrast to the two earlier studies reported in this thesis, the present study adopted a qualitative research design. This required different approaches to the collection and analysis of data, which yielded new and interesting findings and insights. In particular, thematic analysis of the data led to the identification of five themes. In the previous section these themes were described and evidenced in detail. In this section aspects of those themes are discussed in relation to the specific aims of this study, as well as the wider aims of this thesis. The qualitative methods used in this study are also discussed and critiqued in light of the results. The original aims were categorised into four topics: **appropriation**, **meaning**, **impact**, and **design** (see Section 6.1). Consequently, the discussion below is organised according to those topics.

6.6.1 Appropriation

The first aim of this study was to investigate how musicians incorporated the LuminUs into their collaborative music making over the course of four study sessions. Musicians predominantly considered overtly functional and pragmatic uses for the LuminUs; both in practice, and whilst discussing uses. They wanted it to facilitate the communication of clear and interpretable information between them and their co-performer. With respect to the gaze feedback, this mainly concerned the signalling and cueing of musical changes, such as those discussed in the 'aspects of music' theme. The 'expression and communication' theme highlights the roles that glances and eye-contact would normally play in these processes. It appears that these existing functions of gaze inspired the consideration of signalling and cueing as potential uses for the LuminUs. In terms of how this was attempted and achieved in practice, two different approaches were used: a direct approach, used by Group 2; and an indirect approach, used by Group 1. In the direct approach the musician would glance at the extra marker and the light feedback on their co-performer's LuminUs would serve as the signal or cue. In the indirect approach the musician would glance at a marker situated near to their co-performer, and the light feedback would prompt their co-performer to glance back, enabling the signal to be conveyed through mutual eye-contact. In practice, the direct approach was problematic due to the unreliability of the gaze feedback, as evidenced in the 'technology problems' sub-theme. This had less impact upon the indirect approach, since the presence of the gaze feedback was acknowledged by the co-performer's return of gaze. However, Group 1 suggested that they might have opted for a direct approach if the reliability of the device was guaranteed (see Excerpt 51 on page 216). This is an interesting finding, since it implies that musicians would be willing to use technology as a replacement for traditional means of non-verbal communication. Indeed, from an audience perspective, musicians discussed how the absence of visible acts of communication between musicians can be more impressive (see excerpts 8 on page 203; and 9 on page 204).

The inclusion of extra, card-mounted gaze tracking markers was intended to provide additional scope for appropriation of the LuminUs. Two of the groups considered using these markers to monitor how often one of their members looked at a certain object. However, both of these uses were abandoned at the conception stage. In general, the extra markers were either placed near to the musicians as a supplement to the headset markers, or were used instead of the headset markers. There were no instances where the musicians made simultaneous use of the two markers for different purposes. This may have been due to the complexity of dealing with two types of feedback at the same time, especially given that none of the musicians were used to the device.

In contrast to the gaze feedback, there was no consensus on a use for the motion feedback. Group 2 used it as a means of enabling one musician to signal how intense he thought the other's playing should be. However, other than this, none of the groups found a use for the motion feedback.

There was a lot of variation in the time that each group spent experimenting with uses for the LuminUs. Group 1 settled upon a use for the device after the first session and reported increased use of the device over these sessions (see excerpts 37 on page 212; and 38 on page 212). Groups 2, 3, and 4 discussed and experimented with uses for the device during all three dyad sessions. Consequently, it was not possible to observe any specific variations in the use of the LuminUs in relation to whether the groups were rehearsing or composing. The decision to include a composition task also meant that the musicians felt pressured to create something, which diverted their attention away from the technology.

In summary, the appropriation of the LuminUs varied greatly over the four study groups. The only consistently discussed and attempted use of the device was for signalling and cueing using the gaze feedback. However, there were still variations in the functions of the signals and cues, and the approaches to communicating them. The themes identified in this study help to shed light upon these factors. For example, the 'expression and communication' theme provides insight into functions and modes of signalling and cueing, and why certain styles of signalling (e.g. direct or indirect) may be preferable; whilst the 'aspects of music' theme addresses the features of the music that musicians choose to adjust through signalling, such as intensity or volume.

6.6.2 Meaning

This aim focused on the investigation of the ways in which musicians might interpret and draw meaning from the LuminUs feedback. The motion feedback was not widely used by the study groups. However, musicians discussed how they associate expressive movement with the content of the music, such as whether it was calm or intense (see Excerpt 17 on page 206). They also discussed how motion might give them an insight into how a person was "responding to the music" or whether they were "into the music" (see excerpts 7 on page 203; and 19 on page 206). These findings support the results of Study 1, where body motion was found to be significantly correlated with the self-reported engagement and energy.

Potential issues with the motion feedback were that it was not responsive enough, and that it did not represent the right kind of motion. Regarding the latter, a musician pointed out that it would be more useful for the LuminUs to represent the "force" of the movement, rather than its "frequency" (see Excerpt 40 on page 213). This lack of clarity in what the motion feedback was conveying may have led musicians to try to find meaning in the motion feedback by glancing at their co-performer in order to link their actual motions to what was being observed on the LuminUs. Evidence for this was found in Excerpt 41 on page 214. This might partially explain the findings in Study 2, which showed that musicians glanced significantly more when they were receiving motion feedback.

In contrast, the underlying meaning of the gaze feedback seemed to be pretty clear: when the LuminUs lights up it means the other person is looking at one of the markers. For the group who adopted the gaze feedback throughout the sessions, their response to the gaze feedback was to immediately glance towards their co-performer (see Excerpt 42 on page 214). This finding supports the results of Study 2, which showed significantly increased glancing, and a higher proportion of reciprocated glances during the gaze feedback condition. Musicians indicated that glances are normally brief, but that longer glances could carry negative connotations (see Excerpt 13 on page 205). During composition, they also suggested that more frequent glances were seen as an indication to stop and discuss what was being played (see Excerpt 14 on page 205). These findings have a relevance to the findings of Study 1, where the amount of time spent glancing was found to be significantly correlated with self-reported boredom.

The findings of Study 2 showed a significant correlation between the information content (IC) (a measure of musical change) of the pianists' playing, and the number of glances made by their fellow percussionists. This result could be partially explained by a number of findings in the present study. Firstly, there was a very clear association between gaze and the signalling and cueing of musical changes, as discussed within the 'expression and communication' theme. Musicians also discussed looking for appraisal regarding what they were playing (see Excerpt 16 on page 205); and looking to understand and coordinate with what the other musician was playing (see Excerpt 12 on page 205). In both cases, this behaviour could contribute towards a relationship between glancing and musical change.

It is worth noting that the LuminUs may also have served to amplify some of the effects of glance, since it gives the impression that a glance is longer than it actually is, due to the fact that the gaze feedback gradually decreases over a few seconds once a person stops glancing.

Finally, this aim also focused on revealing the wider meanings associated with words such as 'engagement' and 'connectedness'; which were used in the self-report questionnaires in studies 1 and 2. The findings in the present study indicate that musicians do not actually tend to use these words in their discussions of collaborative music making. Instead, they use terms and phrases such as "locked-in" (see Excerpt 46 on page 215), "energy in the room" (see Excerpt 2 on page 202) and "giving yourself into the music" (see Excerpt 20 on page 207). Furthermore, the findings discussed in the 'expression and communication' theme indicate that musicians struggle to provide accurate descriptions of experiential and emotional aspects of collaborative music making.

6.6.3 Impact

The appropriation and meaning aims were concerned with specific aspects of how the LuminUs was used, and the ways in which the feedback was interpreted. In comparison, the impact aim focused upon the wider influence and implications of using the LuminUs. As previously mentioned, there was great variation in the use of the LuminUs across the four study groups. Consequently, the overall impact of the device also varied from group-to-group. The LuminUs's potential for impact appears to be mostly determined by factors relating to, and discussed within the 'context' theme. Specifically, musicians discussed how specific styles, situations, and circumstances surrounding musical collaboration might alter the potential impact of the LuminUs. A common suggestion was that the device would be more useful in larger groups of musicians, or where the spacing between musicians was greater. These suggestions related to the musician's ability to be aware of their co-performers through visual contact. The current study was limited to dyads rehearsing in a small room, meaning that the device's potential for impact was probably not fully exploited. In relation to the gaze feedback, a number of musicians pointed out that they felt the device was of little use, since they were set up to face each other anyway. In order to maximise the usefulness of the device, Group 1 positioned themselves side-by-side, as if they were performing on stage.

Musicians indicated that the LuminUs may be more useful when playing certain styles of music, such as the blues (see Excerpt 77 on page 225). This was predominantly due to the structural flexibility of the musical style, in terms of the presence of cueing and signalling. For the same reason, Group 1 indicated that the device was more useful for composing rather than performing, due to the need to communicate and receive feedback relating to the ongoing composition (see Excerpt 45 on page 215). Furthermore, this group felt that the device had an impact upon their efficiency during the composing sessions (see Excerpt 16 on page 205).

From a more general perspective, one group reported that the experience of using and discussing the LuminUs had an impact upon their awareness of their body motions while playing (see Excerpt 46 on page 215).

6.6.4 Design

This final aim concerned the provision of design considerations and recommendations for the LuminUs, as well as wider implications for the design of affective and social signal-based technologies for collaborative music making. All five of the themes identified in the thematic analysis offer valuable contributions to this aim. The first of these themes is ‘expression and communication’. Within this theme, various functions, styles, and modes of expression and communication are discussed. In its current form, the LuminUs provides simple feedback on *either* motion *or* gaze. However, in reality neither gaze nor motion-based expression ever occur in isolation. Each body movement or glance is always accompanied by additional expressive components, such as head movements and facial expressions. The interpretation of a non-verbal communicative act may be dependent upon observing a combination of these components. For example, when musicians use glances for cueing changes in the music, they will often use a nod of the head to signal the exact moment of the change (see excerpts 9 on page 204; and 18 on page 206). This form of communication initially requires the musician to get the attention of their co-performer(s), and in this sense the gaze feedback of the LuminUs could be useful. However, some musicians discussed using the device as a direct means of signalling and cueing. A potential design modification that could assist in this use would be to incorporate the gaze and motion feedback simultaneously, enabling a signal to be provided in two stages. The first stage would involve gazing at a marker in order to gain the attention of the musician(s), and the second stage would involve nodding to signal the change.

From a motion perspective, it is clear that expressive movements comprise a variety of parameters, such as speed and positioning. One musician discussed how he thought the motion feedback should represent the force of movement, rather than the frequency of movement; giving the example of the LuminUs feedback going “totally red” in response to a musician playing “something soft, very fast” (see Excerpt 40 on page 213). In this case, the motion sensor was positioned around the pianist’s wrist. Clearly, an important design consideration is that movements relating to the musician’s playing actions will vary from instrument-to-instrument, as discussed in the ‘group attributes’ theme. Disassociating expressive motion from functional motion would be a challenging task, since the two are closely intertwined. Indeed, musicians spoke of being able to “hear” aspects of a co-performers emotion in their playing (see Excerpt 20 on page 207), demonstrating how expressive motion transcends through the musician’s instrument. A possible solution would be to focus upon motions that do not tend to be directly involved in playing, regardless of the instrument. Head motions are a good example of this, as highlighted by the musicians in this study (see Excerpt 17 on page 206).

The ‘aspects of music’ theme highlights various musical parameters that the musicians have control over. An implication for design is the way in which musicians relate these parameters to the aforementioned modes of expression and communication. For example, Group 2 chose to use motion feedback to signal the intensity of the music, and gaze feedback to cue structural changes. There are clear relationships here, in that motion and intensity can both change continuously in time, whilst glances and structural changes tend to be more discrete events. The musicians' use of *physical mappings* between expressive behaviours and musical attributes has definite implications for design; especially regarding the ways in which visual feedback is presented. An example of this can be seen in the discussions musicians had concerning whether it was necessary for the gaze feedback to consist of a gradual change in light, instead of just having a single light that went on and off (see excerpts 63 on page 220; and 62 on page 219.

The 'group attributes' theme highlighted the importance of considering individual roles and abilities, and their influence upon group music making. These considerations are also important from a design perspective. The need for design flexibility in accounting for different instruments has already been discussed in relation to motion. However, the type of instrument being played, and an individual's role in the group also determine how useful certain technologies might be to them. For example, musicians pointed out that the LuminUs might be particularly useful for drummers, due to them having limited movement and often being positioned behind other musicians on stage (see Excerpt 72 on page 222). Another group attribute concerns the relationships between musicians. Musicians within some of the groups demonstrated a familiarity with their co-performer's behaviours, such as where they tended to look, or how much they normally moved while playing. This familiarity assisted them with interpreting and understanding their fellow musician's expressions and emotional states whilst playing (see excerpts 20 on page 207; and 30 on page 209). This finding raises the question of whether an affect and behaviour sensing technology, such as the LuminUs, should be designed with the capability to learn about, and become familiar with the affective and behavioural characteristics of individual musicians.

Much like the 'group attributes' theme, the 'context' theme also highlights design issues concerning the versatility of the LuminUs. Musicians pointed out specific situations where they thought the device would be useful, such as in the recording studio, or in an orchestral pit. To what extent should the design of the device be specialised for certain situations? Currently, the LuminUs has only been tested with dyads. For situations involving more than two performers it would be important to consider how the visual feedback could be representative of multiple people. This could simply involve having multiple devices or, as one musician pointed out, using different colours to represent each musician (see Excerpt 65 on page 220).

An important design requirement that was highlighted during this study is the need for the technology to be trustworthy and reliable. If musicians do not have complete confidence in the device then they are likely to avoid using it, especially in performance situations. A related point is that musicians expressed a desire to receive some form of feedback confirming that the other musician had seen their LuminUs feedback – that their 'signal' had been successfully received.

From a more aesthetic and form-related perspective, musicians did not appear to have any problems with the feedback being light-based. Furthermore, they discussed how they were able to observe the feedback in their peripheral vision (see excerpts 33 on page 211; and 37 on page 212). Some musicians reported that the eye-tracking headset was uncomfortable, and that they would not want to wear it for a performance. The latter point was related to audience impressions and presentation, with a number of other musicians also expressing a desire for the technology to be discreet.

6.7 Summary

This study took a qualitative approach to investigating the use of the LuminUs by four established musical dyads. Thematic analysis resulted in the identification of five themes: expression and communication; group attributes; technology; aspects of music; and context. These themes were described in detail and evidenced using excerpts from interviews and discussions between the musicians. The original aims of the study were then discussed in light of the findings across all five themes. A number of these findings were discussed in relation to specific findings from the earlier studies in this thesis. In each case, this provided a broadened insight into affective and behavioural aspects of collaborative music making, as well as the potential applications for sensor technologies.

Chapter 7

Discussion and Implications

This thesis had three principal aims: to investigate how various sensor technologies could be used for measuring affect and behaviour during collaborative music making; to consider applications for these sensors by developing and evaluating a prototype device; and to provide insights into the wider applications for affect and behavioural sensing technologies in the context of collaborative music making. These aims were addressed over the course of three distinct studies. The first of these studies was exploratory in nature, testing an array of different sensors in a highly controlled environment. The results of this study led to the selection of specific sensors and the design of a device – the LuminUs – which was subsequently incorporated into a second, more focused quantitative study. A final, qualitative study, was then undertaken in order to further evaluate the LuminUs and gather subjective data relating to findings obtained in the first two studies. In light of these studies, the work in this thesis can be grouped under three topics:

Sensing musical collaboration: experiences and findings relating to the use of various sensor technologies; their selection, implementation, and informativeness.

Collection and analysis of dyadic data: challenges and solutions relating to the design of dyadic studies, and specific methods and techniques for data collection and analysis.

Design implications: considerations and lessons learnt concerning the design of affect and behaviour sensing technologies for collaborative music making.

This chapter provides a detailed discussion of each of these topics in relation to the work undertaken in this thesis as well as relevant existing research.

7.1 Sensing Musical Collaboration

This thesis investigated a wide variety of sensors for measuring collaborative music making. Figure 7.1 provides an overview of the different types of sensors, their relationships, and the levels at which they can be viewed. As can be seen, sensor data are initially collected from individual musicians, before being combined and processed in order to 'sense' aspects of the musical collaboration. This section introduces basic recommendations for what can be sensed, and how different measures are interrelated (e.g. motion and physiology). Measurable aspects of the musician are separated into external measures (gaze, motion, and musical output), and internal, physiological measures (ECG, EEG, GSR). Each measure is then discussed in detail; covering findings relating to that measure across all three studies, as well as other relevant research.

7.1.1 Sensing the Musician

Musical performance places specific constraints upon the use of sensor technologies. From a practical standpoint, constraints predominantly relate to the physical demands of playing an instrument. For example, sensors that must be attached to the hands, or which restrict body movements, are not suitable for use with musicians. For collaborative performance, practical constraints are also introduced by situational factors. In particular, the positioning of musicians can change over the course of a performance; musicians may not be fully visible because of physical obstructions; and the performance area can be lit by changeable or dim stage lighting. These factors constrain the use of fixed-position, camera-based sensing methods – such as those used for facial expression and motion analysis – since the performance of these methods is degraded

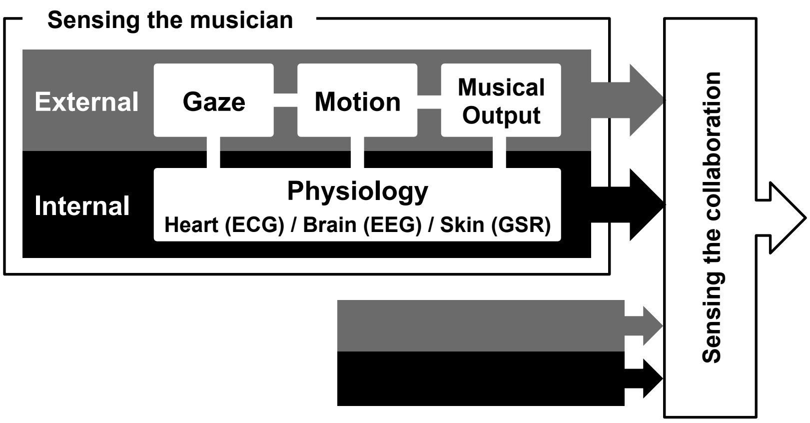


Figure 7.1: Sensing musical collaborations.

by dynamic lighting conditions and occlusions (Lim et al., 2014; Kotsia et al., 2008; Dael et al., 2016). Consequently, the studies in this thesis focused upon body-worn sensors. However, these are also subject to practical constraints. For example, the movements of musicians meant that the EEG and GSR data collected during Study 1 contained substantial levels of noise. In Study 3, movements also caused some problems with the gaze detection whilst using the eye-tracking headsets. For each of the measures discussed in the following sections, specific practical issues and solutions are highlighted.

A less obvious constraint emerged from the findings of Study 3, and it relates to the aesthetics and visual appearance of the sensor technology. Musicians are often conscious of the way that they present themselves in performance situations, and this means that they may be adverse to wearing certain sensor technologies in these situations. For this reason, there may be a preference for sensor technologies that are more discreet.

The sensors that were used in this thesis can be separated according to whether they sense external or internal measures (see Fig. 7.1). Gaze, motion, and musical output are external measures, since they are externally visible to a human beholder. Consequently, features of these measures tend to be directly relevant to expressive and communicative aspects of collaborative music making. As such, it is possible to gain insights into the affective and behavioural phenomena that could be sensed from these measures by simply asking human observers how they interpret them. For example, discussions with musicians during Study 3 revealed that their co-performer's body motions served as an indication of them being 'into the music'. This was reflected by corresponding results in Study 1, which showed correlations between the participants' own self-reported levels of energy and engagement and their body motion, as sensed by an accelerometer.

In contrast, internal measures cannot be directly observed by a co-performer. Within the scope of this thesis, internal measures comprise the physiological measures: GSR, ECG, and EEG. Due to the hidden and subconscious nature of these measures, it is less likely that musicians will be able to provide experiential insights into how they relate to aspects of collaborative music making. Instead, it is necessary to perform quantitative experiments involving the collection of sensor data from musicians. Statistical analysis techniques can then be used to identify relationships between these data and affective and behavioural features of interest. It is important to note that statistically significant relationships alone are not evidence of causal relationships. As such, the researcher must draw upon existing knowledge of psychophysiological processes in order to provide a causal basis to support their findings. This may also involve carrying out further studies. For example, the findings in Study 1 revealed potential relationships between cardiac activity and musical decision making. Explanatory hypotheses were then established based upon existing studies and knowledge, and these hypotheses were tested in Study 2.

It is important to bear in mind that relationships also exist between the different measures. For example, where movement is required to play an instrument, the musician's motion will clearly be related to their musical output. Consequently, musical output will also be related to cardiac activity, since this is directly influenced by levels of physical exertion. This was evidenced in Study 2, where significant correlations were found between the velocity of notes and the musician's HR. Throughout this thesis, measures have predominantly been analysed independently of each other in relation to aspects of collaborative music making. The justification for this is that many of these measures have not previously been investigated in detail in the context of musical collaborations; therefore, it makes sense to begin by studying them in isolation. However, there are numerous findings in this thesis that point towards the potential benefits of combining measures. Where applicable, the following sections discuss ways in which individual measures might be combined with other measures in order to enhance sensor-based analysis of musical interactions.

7.1.2 Gaze

Gaze features were initially extracted through the manual processing of video data from Study 1. Subsequent analysis showed that the percentage gaze time was significantly correlated with self-reported boredom. The importance of visual contact between musicians was also highlighted through the analysis of the effects of the visual condition in Study 1, which indicated that visual contact had a significant influence upon self-reported creativity. These results led to the selection of gaze as a potentially informative measure of musical collaboration.

Existing studies have investigated the functions of gaze in non-verbal communication, and have acknowledged its important roles in non-verbal aspects of musical interactions (see Section 4.2.1). However, it was not possible to find any existing research where automatic methods of sensing gaze had been adopted for the study of gaze interactions between musicians. This was achieved in Study 2, predominantly due to the recent development and availability of affordable, open-source hardware and software for gaze tracking. Recent advances in machine vision also enabled markers to be tracked within a musician's field of vision (FOV). Prior to this, the processing of gaze tracking data generally involved manual, post-hoc analysis of gaze points.

From a practical standpoint, there were a number of issues with the gaze tracking sensors used in studies 2 and 3. The main issues concerned the sensitivity of the marker and pupil tracking to movements and slight changes in positioning respectively. Regarding the tracking of the wearer's pupil, the position of the eye-tracking camera relative to the eye must remain constant. This is achieved through the fact that the headset is fixed on the wearer's head. However, sudden movements or adjustments of the glasses can lead to changes in positioning, which subsequently degrade the tracking accuracy. Regarding the tracking of markers, the movements of the FOV camera, relative to the surrounding environment, can result in motion blur, which makes marker detection difficult. This meant that, in practice, the data from the eye-tracking glasses was not always reliable. This had a negative influence upon participants' adoption of the technology during Study 3. These issues present significant challenges that need to be addressed if eye tacking sensors are going to be successfully used by performing musicians. Having said this, eye-tracking sensors are constantly improving. For example, the current version of the Pupil eye-tracking headset has a FOV camera with a frame rate of 120 frames per second (fps), compared to 30 fps on the headsets used in this thesis[1]. Higher frame rates reduce the level of motion blur (Bellers et al., 2007). A more advanced eye-tracking headset is also available, which boasts wireless connectivity and 'calibrationless setup' (SensoMotoric Instruments (SMI), 2015).

For researchers investigating musical interactions, eye-tracking sensors have the potential to facilitate experiments in which quantitative gaze data can be gathered and analysed. This could help provide empirical support for existing theories surrounding the ways that musicians use gaze; and could also lead to new insights and research questions. For example, the results of Study 2 showed significant correlations between the number of glances made by the drummers and the information content of the pianists' music. This provided quantitative support for the theory that gaze is used as a means of obtaining information (Kleinke, 1986). It also raised a number of research questions, such as – why wasn't the equivalent relationship observed for the glances of pianists and information content of the drummers' playing?

For those interested in developing sensor-based technologies for musicians, gaze tracking clearly has the potential to provide detailed and varied information about real-time collaborative musical interactions. The findings in Study 3 highlighted the importance of gaze in the processes of cueing and signalling between musicians. These findings also indicated that gaze is always used in conjunction with other expressive actions. Therefore, future work could focus on incorporating measures of gaze with

[1] `https://pupil-labs.com/store/`

other measures, such as motion. This could be facilitated by the fact that eye-tracking headsets include a FOV camera. By processing the images from this video stream, additional information such as the motion of the headset, or the motion of objects in the camera view, could be estimated. As face-tracking algorithms improve, it may also be possible to use the video stream to track when a musician is glancing at another musician, without the need for fiducial markers. Furthermore, headsets could be designed to include additional sensors, such as accelerometers, which could accurately track the head movements of the musician.

7.1.3 Motion

For the research conducted in this thesis, basic measures of motion were collected using worn accelerometers. More detailed and accurate methods for motion detection exist; however, they are not well suited to use with musicians in real-world settings. This is predominantly due to the fact that they are camera based, which means that they are adversely affected by occlusions, such as musical instruments. In the case of marker-based systems they also require complex and site-specific apparatus. Despite being one of the simpler methods, accelerometry can still be used to extract detailed information about features of motion. A quantity of motion (QoM) feature was used in the studies reported in this thesis, since previous studies have shown this to be an informative feature for the detection of affective states (see Section 4.2.2).

Accelerometer placement is an important consideration, from both practical and feature-processing standpoints. In Study 1 accelerometers were positioned on the musicians' heads, bodies, and wrists. The quantity of body motion was found to be positively correlated with self-reported creativity, engagement, and energy. The specificity of these results to body movements is evidence of the fact that certain placement sites are more informative than others. The results are also a promising indication of the potential uses for accelerometer-based motion sensing in the context of musical collaboration. However, the fact that multiple self-report items were all correlated with the same motion feature suggests that basic features of motion may not be discriminatory enough to classify distinct forms of expression and their underlying affective meanings. If real-world applications are to be created then it may be necessary to test and develop more advanced methods for processing and interpreting accelerometer data. This could involve incorporating data from multiple sites (e.g. body and head), and using machine learning methods in order to model and detect specific expressive motions. It may also be necessary for the data processing to be customised to specific individuals or instrumentalists according to their specific styles of movement. The findings in

Study 3 highlighted how musicians become familiar with aspects of their co-performer's movement behaviour. Motion sensor technologies may also need to exhibit a similar capacity for learning.

As previously discussed, musical output, physiological measures, and gaze behaviour are all directly influenced by types of motion. This was evidenced in the results of Study 2, and is illustrated in Fig. 7.1. This factor, coupled with the ease of collecting accelerometer data, forms a strong case for the inclusion of motion sensors in any system designed to measure affect and behaviour during collaborative music making. For example, motion data could be used to account for the influence of physical exertion upon cardiac activity; enabling the extraction of cardiac features that were more representative of psychological processes. A similar approach has already been adopted by Dean and Bailes (2013), who used a motion sensor to account for the influence of movement on skin conductance readings taken from the left ankles of improvising pianists. In each case the motion sensor and skin conductance sensor were attached to the same ankle. Subsequent time series analyses resulted in the identification of relationships between skin conductance changes and transitions between musical segments, but only after the impact of movement had been accounted for.

7.1.4 Musical Output

During the studies in this thesis, MIDI and audio data were used for the extraction of features relating to the musical output of individual musicians. On the surface, these data might not seem to be directly related to the affective and behavioural phenomena with which this research is concerned. However, musical output contains important information relating to the expressive features of musical actions, such as timing and dynamics. In the case of improvised music, it also represents the musical decision making processes of the musician. This was the main justification for extracting musical output features in this thesis; given its focus upon collaborative improvisation. In Study 1, MIDI data were used to analyse the timing synchrony between the dyads. MIDI and audio data were also analysed manually in order to extract rhythmic change points (RCPs). The increased size of Study 2 meant that automated methods were sought for extracting similar features. Two features were used: information content (IC), and musical change points (CPs). The former was computed using existing software, whilst the latter was extracted using software specifically developed for the study (see Section 5.3.4). Both features were intended to represent the timing and overall quantity of structural changes in the music. Subsequent analyses showed that both features provided valuable insights into the musical output, and the way that it related to other quantitative measures. These findings varied between the two features, and were also instrument-specific. It was noted that IC and CP only provide objective representations of the musical output. In the study of musical collaboration, qualitative measures of musical output are also important, especially when evaluating aesthetic outcomes, such as creativity. Due to their inherently subjective nature, the accurate extraction of such measures through automated means is an extremely difficult task (Galanter, 2013).

The aforementioned approach to the extraction of musical output features was inspired by existing work in music information retrieval (MIR): a growing field of research, which concerns the computerised extraction of information from music (Downie, 2003). MIR has a wide array of uses, including the identification of tracks by specific composers, and the classification of music according to concepts surrounding mood and emotion. Casey et al. (2008) discuss high-level descriptions of music that can be extracted, such as timbre, pitch, harmony, and structure; as well as low-level audio features, such as spectral properties of the audio signal. In the context of collaborative music making, MIR methods can provide researchers with tools for performing quantitative analysis of musical outcomes. This affords an extra layer of insight into the musical interactions, which can subsequently be analysed alongside features obtained from other quantitative measures. However, MIR methods predominantly utilise pre-recorded music, represented in either digital-audio or symbolic formats (e.g. MIDI or musical score). Consequently, they are somewhat restricted to post-hoc analysis of musical output. Furthermore, the capture and representation of the musical output are dependent upon numerous factors, such as the instruments being played and the performance situation. The studies in this thesis utilised MIDI instruments in order to record symbolic representations of individual musicians' playing. Most acoustic instruments do not have MIDI capability, and would need to be represented as audio. This introduces additional technical issues such as source separation - the isolation of individual instruments within a single audio recording.

7.1.5 Physiology

In contrast to the external measures discussed in the previous sections, physiology is classed as an internal measure due to the fact that it cannot be directly viewed by an external observer. Having said this, some *indicators* of physiological functioning may be externally visible, such as perspiration, or changes in skin colour due to blood flow. However, these cannot provide detailed insights into the underlying physiological processes.

Physiological activity tends to function outside of conscious awareness and control. This means that physiological measures can reveal aspects of a person's behavioural and affective states that might not be expressed through actions and words. This is especially useful for studies of collaborative musical performance, where actions are restricted and verbal interaction is scarce.

All of the physiological measures used in this thesis involve the use of electrical sensors attached to the skin. In most cases, the raw data collected from these sensors must undergo a degree of processing before meaningful features can be extracted. The complexity and quantity of processing varies from measure-to-measure; as does the susceptibility to noise. These factors must be taken into account when selecting physiological measures for specific studies or applications. In Study 1, three physiological measures were collected: EEG (brain activity), ECG (cardiac activity), and GSR (perspiration). Features extracted from all three measures showed interesting correlations with self-reported aspects of the musicians' collaborations. However, the decision was made to exclude EEG and GSR from further studies due to their sensitivity to motion artifacts.

In studies 1 and 2, relationships between creative decision making and cardiac activity were investigated. The results from Study 1 showed relationships between HR extrema and decision points. However, these were not observed in all participants. In Section 4.2.3 a comprehensive review of existing research relating to cardiac activity and decision making was undertaken. This informed the development of specific hypotheses, which were subsequently tested in Study 2. The findings from this second study were inconclusive; showing limited evidence for relationships between features of cardiac activity and decision making (see Section 5.6.2). The lack of clear correlations was attributed to the fact that cardiac activity is influenced by an array of different processes. Therefore, in order to specifically identify cardiac changes that correspond to a sub-set of these processes, it would be necessary to account for the extraneous processes in some way. As previously discussed (see Section 7.1.3), researchers have already attempted to achieve this for physiological measures from pianists; using motion sensors to model and account for the influence of movement upon skin conductance data. In this case, the researchers were accounting for motion artifacts in the skin conductance readings, rather than the influence of motion on perspiration levels. Modelling and accounting for the *internal processes* that influence a physiological response is a far more complex problem, due to the fact that physiological functioning varies between subjects (Cacioppo et al., 2007a), and can be influenced by a wide array of factors (Blascovich and Mendes, 2010; Bradley and Lang, 2007). With regards to cardiac activity, these could include physical factors, such as the subject's age and fitness levels; chemical factors, such as the use of medication; environmental factors, such as the temperature; and psychological factors, such as stress (Berntson et al., 2007). In the context of musical interactions, the factors of greatest interest are likely to be those that have a transient, short-term influence upon cardiac activity. For example, the presence of an audience might influence a musician's anxiety levels; whilst the venue environment might influence their body temperature. Accounting for these factors will require parameters to be specified or learnt at the level of the individual; since the nature of these influences will vary from person-to-person.

7.2 Collection and Analysis of Dyadic Data

The three studies reported in this thesis adopted a variety of approaches to the collection and analysis of dyadic data. These included qualitative and quantitative designs; naturalistic and controlled settings; physiological, motion, gaze, musical performance, self-report, interview and observational data sources; and individual, dyadic and group-level analyses. Some of these approaches, such as the use of linear mixed models, were directly adopted from existing research methods; whilst others, such as the extraction of musical decision making features, were developed and modified in order to suit specific requirements of the studies. In this section the considerations, challenges, and issues associated with carrying out dyadic studies are discussed in detail. These discussions are contextualised with respect to collaborative music making; however, they also have relevance to dyadic studies in general. The section is organised according to the chronological stages of carrying out a study, namely: study design, data collection, and methods of analysis.

7.2.1 Study Designs

Collaborative music making incorporates an extremely broad range of activities, which take place in a variety of settings, with differing musical styles, and a spectrum of musical abilities. Consequently, it is not possible for any single study to investigate all aspects of collaborative music making. Instead, it is necessary to focus upon a restricted sub-set, which will be determined by the specific aims and objectives of the research. The breadth of the study will also have a direct impact upon the generalisability of the results and findings. The research in this thesis focused specifically upon collaborative improvisation within experienced dyads. In the first two studies the musical instruments were also restricted, and participants were given specific improvisation tasks.

This means that the findings in this thesis must be considered in light of these restrictions. Having said this, it is possible that some of the findings could be generalised to other collaborative activities, both musical and non-musical. Further research would be required in order to verify the generalisability of these results.

The varied nature of collaborative music making is concomitant with the fact that it is highly situated and context-dependent. This also means that it is not well suited to study under controlled, laboratory settings. Unfortunately, when working with quantitative sensor data, controlled study designs are a pre-requisite to performing valid statistical analyses. Addressing this conflict between ecological and statistical validity was one of the biggest challenges when designing the studies in this thesis. Where possible, the experimental settings were designed to resemble real-world musical performance spaces; using stage lighting and purpose-built performance spaces. However, the participants were still aware that they were taking part in an experiment, which could have had an impact on their behaviour. The effectiveness of simulated musical performance settings has been investigated by Williamon et al. (2014), who developed and tested a virtual 'performance simulator' for training musicians at the Royal College of Music, London. The simulator consisted of a small performance space with a screen, onto which a virtual audience or audition panel were projected. Curtains and spotlights were positioned either side of the screen, and loudspeakers provided noise distractions, such as coughing and phone-ringing. A small backstage area was also created, with mock CCTV footage of audience members sitting in the performance space. When testing the simulator with violinists, participants' levels of reported anxiety and patterns of heart rate variability (HRV) were found to be comparable between simulated and real performance situations. These results are a promising indication of how virtual environments could facilitate the design of experiments that are both ecologically and statistically valid.

Another major consideration when designing experiments is whether the experiment should be hypothesis-driven (confirmatory) or exploratory. Exploratory studies are appropriate when little is known about the area of interest. In the social sciences, exploratory research has been promoted as a means of avoiding the formulaic restrictions of confirmatory research (Stebbins, 2001), and revealing previously unimagined connections and casual mechanisms (Reiter, 2013). It was for similar reasons that an exploratory approach was adopted for the first study in this thesis. This successfully resulted in interesting and unforeseen findings, which were then used to inform the design of a more focused, confirmatory second study.

The third study in this thesis was also exploratory; taking a qualitative approach, which involved the thematic analysis of interview data. The main purpose of the third study was to evaluate musicians' use of the LuminUs. However, this also resulted in unforeseen findings that could have served as a basis for further research. Indeed, qualitative research has traditionally been used as a precursor to quantitative studies; enabling areas of interest to be identified (Ritchie, 2003). Having said this, it has been noted that "one of the most underutilised ways of using qualitative and quantitative research together is to follow statistical research enquiry with a qualitative study, yet this is a particularly powerful way of combining the two approaches" (Ritchie, 2003, p. 42). This was the approach taken in this thesis; justified by the fact that the research initially set out to explore potential uses for various types of sensor data; and that these uses were best explored through statistical enquiry, due to the quantitative nature of the data. Building upon the quantitative findings of Study 2 with a qualitative study meant that it was possible to frame particular statistical results in the context of more subjective and qualitative aspects of collaborative music making.

7.2.2 Data Collection Methods

The studies in this thesis adopted both quantitative and qualitative methods of data collection. Each of these presents particular challenges and issues when conducting studies with collaborating musicians. With regard to quantitative methods, multiple factors must be taken into account when selecting which data to collect, and which sensors to use. The different measures that were used in Study 1 were evaluated according to their reliability, informativeness, and practicality (see Section 3.5.3):

Reliability concerns whether the data collected from a particular sensor can be trusted as a representation of the intended measure. The main issues affecting reliability are the accuracy of the chosen sensor, and its susceptibility to noise. It is also important to take into account the sampling rate at which the data are collected. In Study 1 ECG data were collected at 51.2 Hz, which was not a sufficient sampling rate for accurate HRV analysis to be performed. Consequently, the sampling rate was increased to 512 Hz for the collection of ECG data during Study 2. The selected sampling rate is usually dependent upon the frequency of change in the signal being measured. For example, changes in perspiration are much more gradual than changes in the electrical activity of the heart; therefore, sampling rates for skin conductance data do not need to be as high as those for ECG data.

Informativeness concerns the amount of useful information that can be extracted from the data. In the case of Study 1, the informativeness of each measure was not known prior to the study; and one of the study aims was to evaluate the informativeness of different measures. Informativeness is very much dependent upon the ways in which the collected sensor data are subsequently processed and analysed. This is discussed below.

Practicality is discussed earlier in this chapter (see Section 7.1.1).

When collecting data from multiple participants or sensors, it is important to ensure that all the data can be synchronised to a common timeline. The required precision and accuracy of this synchronisation is dependent upon the type of data being collected and the objectives of the research. For example, if the research is concerned with relationships between features that have been extracted and averaged over time periods in the order of minutes, then it may not be necessary to achieve synchronisation accurate to the millisecond. However, analyses of discrete time events may require much higher accuracy. One of the challenges for achieving synchronisation is that individual sensors often have proprietary software; meaning that data must be saved within multiple different programs, each with their own time references. In Study 1 a novel method was used for synchronising the data sources manually. This involved creating a series of short time-duration events that were common to all data sources, and which could subsequently be used to align the different data sets. In the second study, data from multiple sources were transmitted to a single program, enabling them to be saved against a shared time reference. This was facilitated by the fact that the selected sensors included open source software, which could be modified in order to transmit the data to a single program over a shared network. It is important to note, however, that saving data from multiple sources against a common timeline does not guarantee accurate synchronisation. This was highlighted by the fact that the MIDI data collected in Study 2 were found to contain a time drift. Timing inaccuracies can also be introduced by delays in the transmission of data from the sensor to the software program. This can be a function of the speed and bandwidth of wired and wireless networks used to transmit the data. It is important for researchers to be aware of these inaccuracies and to take them into account when collecting and analysing data. This should involve carrying out tests in order to determine the extent of timing inaccuracies prior to undertaking experiments.

Quantitative self-report data were also collected during the first two studies in this thesis. In both cases, this involved the use of scale-based, structured questionnaires. One of the issues with developing and evaluating questionnaires for collaborative music making is that it is difficult to perform repeated tests. When the musicians are improvising – as they were in studies 1 and 2 – this becomes especially challenging, due to the unique nature of each improvisation. Consequently, it may not be possible to assess the validity and reliability of custom questionnaire measures. This can be addressed by adopting and adapting existing questionnaires; a method that was partially used to develop the questionnaires in this thesis. However, the fact that these questionnaires may have been developed and tested for different activities and purposes means that their validity and reliability may not be applicable in the context of collaborative music making.

With regards to the collection of qualitative data, semi-structured interviews and observations were used in the third study in this thesis. In this case, the aim of the study was to evaluate the musicians' use of the LuminUs, rather than to analyse sensor data. When conducting interviews, it is necessary to be aware of the influence that the interviewer can have upon the data collection. This is especially relevant when the interviewer is also one of the researchers, since there is a risk of them revealing aspects of their hypotheses and preconceptions concerning the outcomes of the experiment. Orne (1962) describes how cues that convey aspects of experimental hypotheses to participants can significantly impact the participants' behaviours. He refers to these cues as 'demand characteristics'. Demand characteristics have been shown to lead participants to respond in a way that will support the experimental hypotheses, and are closely related to wider issues concerning experimenter bias (Klein et al., 2012). During Study 3 participants were aware that the interviewer had been involved in the development of the LuminUs, due to the fact that many of them had also participated in the second study. Consequently, they may have responded more positively to questions about the LuminUs in order to please the researcher. This issue could have been avoided to some extent through the use of an independent interviewer. However, this may have led to interview data lacking in relevant depth; due to the need for the interviewer to have a detailed knowledge of the research in order to pursue appropriate topics during the interviews.

7.2.3 Analysis Methods

A variety of analysis methods are reported throughout this thesis, owing to the breadth of both quantitative and qualitative data collected, and the use of different levels of analysis: individual, dyad, and group. A common starting point for the analysis of both qualitative and quantitative data was the use of visualisations. These were especially useful where the approach to the analyses was exploratory, as in studies 1 and 3. In Study 1, strip-plot visualisations were created to represent multiple features along a common timeline (see Fig. 3.5 and Appendix A.3). Coloured segments represented periods of gaze by each participant in a dyad; shaded areas represented continuous heart rate measures; and discrete markers represented rhythmic change points. In Study 3, thematic maps were created, enabling the visualisation of links between different codes and themes and their prevalence in the data set. In both cases, these visualisations aided the early stages of data analysis; highlighting potentially interesting relationships between different measures. The processing and visualisation of large and complex data sets is facilitated by ongoing advances in computing; providing researchers with valuable tools for exploring and analysing their data through visual inspection. These advances can be further exploited by using the word wide web as a means of crowdsourcing volunteers to assist in the visual analysis of large data sets. This distributed data analysis has been used to great success for a range of research objectives, including the identification of new galaxies (Simpson et al., 2014), and the matching of whale calls using spectrograms and audio samples (Sayigh et al., 2013).

An important factor to take into account when analysing dyadic data is non-independence (Kenny et al., 2006); which refers to the fact that data from individuals within a dyad may be related. Kenny et al. (2006) describe various types of 'linkage' that can occur between members of a dyad, two of which are relevant to the studies reported in this thesis: *voluntary linkage*, and *experimental linkage*. The former refers to linkage that occurs between friends or couples, and which normally develops over time; whilst the latter refers to relationships that are created in the laboratory, when two people are asked to interact. Non-independence may occur as a result of a combination of these two types of linkages. Additionally, Kenny et al. (2006) describe four sources of non-independence:

Compositional effects: These effects are the result of dyad members sharing similarities prior to being grouped together in a study, and will often occur when participants are grouped in a non-random way.

Partner effects: These occur when characteristics or behaviours of one member of a dyad influence the other member's data.

Mutual influence: This is a process of feedback, whereby the outcomes from both members of a dyad directly affect one another.

Common fate effects: These occur when both dyad members are exposed to the same causal influences.

All four of these sources could lead to non-independence in data collected from pairs of collaborating musicians. For example, compositional effects could emerge due to musicians in an established dyad having similar musical tastes and backgrounds; partner effects and mutual influence could be the result of the behaviours and musical contributions made by each dyad member during a performance; and common fate effects could be the result of a shared stimulus, such as the animation used in Study 2. Sources of non-independence can be mitigated by the design of a study; especially those that occur due to voluntary linkage and compositional effects. This was the main reason why participants in Study 2 were selected such that dyad members did not know each other and had not previously played together. However, sources of non-independence that arise due to experimental linkage must be taken into account during the data analysis. Standard statistical analysis methods, such as analysis of variance (ANOVA) and multiple regression, make the assumption of independence between individual data points. Therefore, unless independence can be determined, these methods are normally unsuitable for dyadic data analysis (Kenny et al., 2006). Consequently, in the statistical analyses reported in this thesis, tests were performed to establish the independence of dyadic variables, and appropriate statistical methods were chosen to deal with non-independence.

One of the most prevalent statistical methods used in this thesis was linear mixed modelling. Linear mixed models (LMMs) facilitate the analysis of data that contains repeated measurements from the same statistical units. Therefore, they are suitable for cases where measures are not independent. LMMs are also useful when the data are split over various grouping levels, such as individual, dyad, and study group.

The importance of considering data analysis over various levels is highlighted in the discussion of the findings in Study 1 (see Section 3.5.2), and is also evidenced by the findings in Studies 2 and 3. In both cases, some results were generalisable across dyads, whilst others seemed to be specific to particular dyads, and to the individual members of the dyads. This is reflective of the nature of music making, whereby certain characteristics, such as the capabilities of an instrument, are common to all musicians playing that instrument; whilst more subjective characteristics, such as playing style and expression, are specific to individuals. These findings support the need for multiple levels of analysis when investigating collaborative music making.

The quantitative analyses in this thesis were either performed using discrete features, or using continuous time series features. Whilst the former is amenable to analysis with common statistical techniques, such as LMMs and t-tests; the analysis of correlations between time series presents greater challenges. In particular, behavioural and physiological signals tend to exhibit autocorrelation, which must be accounted for prior to analysing cross-correlations between time series. Delays are also likely to exist between these signals and their associated stimuli; meaning that time-lagged correlations should be taken into account. Furthermore, when averaging over multiple time series to create a single, averaged time series, it may be necessary to normalise individual series prior to averaging. These issues are discussed further in sections 5.5 and 5.6.2, in relation to the time series analysis of cardiac and musical decision making features. There are a wide variety of approaches to time series analysis and a lack of consensus on appropriate methods. Consequently, researchers must have a good understanding of the nature of their data and the objectives of their analyses, in order to select a suitable approach.

A qualitative method of analysis was used in the final study in this thesis, as a means of evaluating the LuminUs. As with many qualitative analysis methods, this involved manually transcribing and coding data in order to identify and describe themes within the data set. One of the main issues with these methods is that the manual processing of data can be time-consuming and subject to experimenter bias. As automated speech and language analysis tools improve, it is possible that some of these processes could become automated in the future.

7.3 Design Implications

This section discusses implications for the design of affect and behaviour sensing technologies in the context of collaborative music making. The section initially considers the potential applications for these technologies, before discussing three specific aspects of technology design: sensor selection, signal processing, and visual feedback. Findings and examples from the studies involving the LuminUs are used alongside examples from existing studies and technologies, in order to address wider issues surrounding technology design for musicians.

7.3.1 Potential Applications

From a design perspective, the original aims of this thesis were to contribute towards the development of technologies that incorporate affect and behaviour sensing in order to enhance non-verbal interactions between collaborating musicians. In the wider context of affective computing and social signal processing, there has been much research on the development of sensor-based technologies for sensing human affect and social interaction. However, at the time of writing, there are very few examples of these technologies being adopted in real-world applications. A potential reason for this is that research is often driven by a desire to explore and demonstrate the capabilities of the technology, rather than address the specific demands and needs of technology users. The research presented in this thesis was jointly motivated by a desire to explore technological capabilities, and a desire to focus on the needs of a specific sub-set of technology users - collaborating musicians. However, the approach to the research was still somewhat technology oriented, since the first study in the thesis was designed primarily to assess the capabilities of various sensors. Having selected certain sensors, a prototype device was designed and evaluated in relation to its potential uses and impact upon collaborating musicians.

The outcomes of Study 2 indicated that providing simple visual feedback on the glances of a co-performer had a significant influence on the glancing behaviours of musicians. These findings suggest that technology has the potential to influence non-verbal interactions between collaborating musicians. However, this study did little to evaluate the extent to which this influence might be beneficial to musicians and the outcomes of their collaborations. These issues were addressed in Study 3, where specific situations and circumstances were identified, in which musicians envisaged potential applications for the technology. Two of the most prevalent use-cases that emerged from the analyses were orchestral and recording studio performances. These situations share common features. Firstly, they often involve physical separation between musicians, which reduces the potential for non-verbal interaction. In the case of orchestral performance this might be due to the fact that musicians can be spaced far apart, or that they often play in an orchestral 'pit', where lighting is limited. For studio performances musicians can be separated into different rooms or booths whilst they record their parts. Secondly, both situations can involve a heightened need to be attentive to distinct musical cues. For example, members of an orchestra must take note of cues provided by the conductor; whilst, in a recording studio, musicians may need to play to a click track. In these situations, technology has the potential to give musicians an improved sense of the behaviours, non-verbal signals, and emotions of their co-performers.

The LuminUs processes sensor data in real-time and uses it to provide continuous feedback to musicians during performance. An alternative way of using sensor technologies would be to collect and store data over the course of a performance and then allow musicians to review and interpret these data at some point after the performance. For example, a musician in Study 3 expressed an interest in knowing exactly how much they moved during performances. The use of sensor technologies for self-measurement and tracking has become increasingly popular over the last decade; a trend encompassed by the popularised term, 'Quantified Self' (Swan, 2013). This trend has been accelerated by the development of small, mobile and affordable technologies for sensing and recording data. In particular, these developments are closely related to the emergence of wearable technologies, such as smartwatches and activity trackers (Russo, 2015). A number of researchers have investigated the use of wearable technologies for measuring aspects of musical performance. For example, Benning et al. (2007) used various types of motion sensor to capture body gestures and train a student to play the tabla; Dalgleish and Spencer (2015) developed a wearable system for helping trumpet players maintain good posture, using a 3D camera to record their posture, and wearable vibrotactile arrays to provide feedback during playing; and a wearable system involving ECG and accelerometer sensors was developed by Kusserow et al. (2010) to measure stage fright in musicians. Only a single example was found of wearable sensors being used in the context of collaborative musical performance: Gloor et al. (2013) investigated group flow amongst jazz musicians, using 'sociometric badges' containing accelerometers and microphones to measure and compare the 'energy levels' of the performers.

Consumer applications for wearable technologies have predominately focused on the health and fitness markets. At the time of writing, the Soundbrenner Pulse was being marketed as the World's first wearable device for musicians (Soundbrenner, 2016). This wrist or ankle worn device simply acts as a metronome by vibrating at a fixed rhythm set by the user. In the coming years it is envisaged that more wearable technologies targeted at musicians will emerge.

7.3.2 Sensor Selection

As previously discussed (see Section 7.2.2), three parameters were used to evaluate the sensors used in the studies in this thesis: practicality, reliability, and informativeness. Whilst informativeness is highly dependent upon the way in which the sensor data are processed, all three of these parameters should be taken into account when selecting sensors to incorporate into the design of new technologies. An additional parameter, which is specifically relevant to commercial technology development, is *compatibility*. Compatibility encompasses the extent to which the type of sensor is able to work in conjunction with existing technologies and devices. For example, if it is possible to connect the sensor to a smart-phone, then an app on the phone could perform the data processing and provide visual feedback. Alternatively, the app might utilise the sensors already built into the smart-phone, such as accelerometers and cameras. Consumer technologies, such as smart phones and watches, are being developed with an increasing array of in-built sensors, many of which could be utilised for the measurement of performing musicians.

Compatibility also concerns the way in which the sensor communicates its data. For example, it might use Bluetooth, which is a popular protocol for wirelessly transmitting data between electronic devices. There are also protocols that are specifically relevant for musical performance, such as the MIDI protocol, used to communicate with electronic instruments; or the DMX protocol, used to control stage lighting.

It is possible that sensors could be integrated into the instruments themselves. For example, a wireless accelerometer could be incorporated into a violin bow in order to measure aspects of the musician's expressive playing. However, this would clearly restrict the use of the technology to certain instruments, and could also lead to increased costs. Having said this, new interfaces for musical expression often incorporate the kinds of sensors studied in this thesis (Medeiros and Wanderley, 2014). These instruments offer the potential for affective and behavioural data to be collected and processed without the requirement for additional hardware.

7.3.3 Data Processing

As discussed in the previous section, the informativeness of any given sensor is highly dependent upon the ways in which its data are processed. Data can either be processed in real-time, or delayed-time. In order to provide sensor-related feedback to musicians *during* a performance, the processing of data must be near real-time. The speed at which data can be processed is dependent upon the computer assigned to perform the processing, the number of calculations required to process each data sample, and the sampling rate of the sensor. In the case of the LuminUs, the speed at which image frames are captured and processed from the FOV camera directly impacts the levels of motion blur and subsequently, the ability to detect the gaze markers. The calculation of some of the features used in post-hoc analyses, such as those relating to heart rate variability (HRV), requires multiple measurements taken over a particular time frame. If these features were used in a real-time system then the computer would need to continuously store and update a set of data values in order to perform the required calculations.

In the context of collaborative music making, it may also be necessary for data to be combined and processed from sensor measurements relating to multiple musicians. Increasing the number of data streams that must be processed also has a direct impact upon the number of calculations that must be performed. Therefore, for large groups of musicians, such as orchestras, it may not be possible to process data from every musician in near real-time using a single computing device. An alternative option would be to use 'cloud' computing, whereby the data are uploaded to a network and processed using powerful cloud-based computers. However, this would mean that access to the internet would be required in order for the sensor to be used, which would be prohibitive for musicians playing in venues that lacked such resources.

Results throughout the three studies in this thesis highlighted the importance of taking into account individualistic aspects of the musician when interpreting sensor data. Consequently, it may also be necessary to tailor the data processing according to individuals. This could be achieved by asking the musician to provide specific details relating to attributes such as their personality, behaviours, or style of playing. It could also be achieved by collecting and storing baseline data from the musician, enabling the data processing to be adjusted accordingly. An additional option would be to use machine learning techniques in order to 'tune' the processing of sensor data to the musician over time. This could involve either supervised, or unsupervised learning. For the former, the musician might provide the device with subjective feedback following each music making session; allowing the device to learn how subjective ratings related to the sensor data. In the case of unsupervised learning, the device could attempt to learn how features derived from multiple sensors were related. For example, it might use MIDI data to extract musical change features and then learn how these related to cardiac activity features.

An important aspect of the data processing is that its output must consist of information that can be meaningfully interpreted by the user. For post-performance analysis of the data this might involve the use of complex graphical and numerical representations. Whereas, for live performance feedback, simpler visual representations might be required, such as those provided by the LuminUs. The provision of visual feedback is discussed further in the following section.

7.3.4 Visual Feedback

The provision of meaningful feedback is an essential element in the design of any sensor-based technology for collaborating musicians. This section focuses upon a discussion of *visual* feedback, since this was the method adopted in the design of the LuminUs. More specifically, the LuminUs used a strip of coloured LED lights to provide simple and dynamic feedback in a way that was not excessively distracting for the musicians.

A musician's ability to attend to visual feedback is somewhat dependent upon the situation, context, and nature of their playing. For example, one of the drummers in Study 3 adjusted the positioning of their LuminUs according to where they were playing on the drum kit. For musicians playing in live situations, additional visual distractions, such as the audience or stage lighting, might also influence their visual attention. These issues predominately concern physical aspects of the visual feedback, such as its positioning and visibility. Consequently, devices designed for wide-ranging use should afford musicians a degree of control over aspects of the feedback, such as positioning and brightness.

The use of simple, light-based feedback was sufficient to influence the behaviours of the musicians during Study 2. Furthermore, musicians in Study 3 discussed how they were able to attend to the LuminUs in their peripheral vision. Through the use of both space and colour, the lighting used by the LuminUs has the potential to represent dynamic and multidimensional information. In this case, different colours were used to represent different types of feedback (gaze or motion), and the number of illuminated lights represented the length of the gaze, or quantity of motion. To ensure that the data are represented in a simple and efficient form, the dimensionality of the visual display should be no greater than the dimensionality of the data to be represented. A related issue arose during discussions with musicians in Study 3, where it was suggested that the gaze feedback should simply consist of an on/off light; since the only important information was whether the musician was gazing, or not gazing at the marker.

The colour combinations used for the LuminUs feedback were influenced by those used in existing displays, reviewed in Section 4.2.4. Musicians in Study 3 reported that they felt that the colour selections were suitable for the types of feedback. Research has shown that chroma (colour intensity) is positively related to the three affective dimensions: arousal, valence, and dominance (Suk and Irtel, 2010). However, associations between colour and emotion have also been shown to be dependent upon personal preference and past experiences (Kaya and Epps, 2004). There is very little research on the affective and behavioural interpretations of colour in the context of musical performance.

7.4 Summary

In this chapter the work undertaken throughout this thesis has been drawn together in order to highlight prominent findings and issues relating to the use of sensor technologies for measuring and enhancing affective and behavioural interaction during collaborative music making. The discussion was split into three sections, covering the use and selection of sensors; the collection and analysis of dyadic data; and the implications for the design of real-world devices. These topics have relevance to those working with affect and behaviour sensors across a range of fields, from academic research to commercial technology design and development. Furthermore, this is not restricted to the context of collaborative music making. The use of sensors for measuring affective and behavioural phenomena has potential relevance within the wider scope of the performing arts, and beyond. For example, sensors have been used to investigate theatrical dyadic interactions (Metallinou et al., 2013); expressive gestures in dance performances (Camurri et al., 2003); interactions with new media art (Morgan and Gunes, 2013); and the detection of conversing groups (Hung et al., 2014).

Chapter 8

Conclusion

The previous chapter brought together the work undertaken throughout this thesis and discussed it according to three topics: sensing musical collaboration; the collection and analysis of dyadic data; and design implications. These topics are a testament to the breadth and multidisciplinary nature of this thesis, which incorporates exploratory research; qualitative and quantitative studies; and technology design, development, and testing. This final chapter provides a concise overview of the work undertaken, and highlights some associated limitations. It concludes by looking ahead, placing this research in a wider context and imagining future work.

8.1 Overview and Major Findings

The work in this thesis comprised an exploratory study (Chapter 3); which led to the design and development of a prototype device - the LuminUs (Chapter 4); which was subsequently tested in a quantitative/controlled study (Chapter 5) and a qualitative/longitudinal study (Chapter 6). This section provides a concise overview of the work undertaken, and the major findings reported in each of these chapters.

8.1.1 Study 1: Sensing collaborating musicians (Chapter 3)

This research began with an exploratory study, which was designed to address the first aim of this thesis:

A1 Investigate how best to continuously measure behaviour and affect during dyadic musical interactions.

More specifically, the study aimed to assess a range of sensors; identify informative measures and features; and report exploratory findings. Five pairs of experienced percussionists took part in the study, which involved playing basic improvised rhythms on a drum pad under two conditions: one where the dyads were visible to one another, and the other where they were separated by a screen. Eight quantitative measures were collected from each percussionist during their performances. These included three physiological measures captured using ECG, EEG, and GSR sensors; three motion measures recorded from accelerometers worn on the wrist, body, and head; and performance and behavioural measures collected using video cameras and MIDI data. Self-report data were also collected using a post-performance questionnaire.

Data from the eight measures were processed in order to extract 26 different quantitative features for each participant, under each condition. These features were then analysed at individual, dyad, and group level, using visualisations and statistical tests. The analyses revealed interesting relationships between musical decisions and changes in heart rate. Self-reported measures of creativity, engagement, and energy were correlated with the quantity of body motion; whilst EEG beta-band activity was correlated with self-reported positivity and leadership. Regarding co-visibility, lack of visual contact between musicians had a negative influence on self-reported creativity. The number of glances between musicians was positively correlated with rhythmic synchrony, and the average length of glances was correlated with self-reported boredom. The results indicated that ECG, motion, and glance measurements could be particularly suitable for understanding and enhancing affective and behavioural interaction between collaborating musicians.

8.1.2 The LuminUs: Designing a device for collaborating musicians (Chapter 4)

This chapter documented the design and development of an affect and behaviour sensing device to enhance collaborative music making. In light of the findings from the first study, the following input measures were selected: cardiac activity, gaze, and motion. An additional design consideration was the output modality for the device. This had to be capable of representing the processed input measures, whilst also satisfying the user and contextual requirements associated with collaborative musical performance. In view of these requirements, it was decided that the output modality should be visual, minimal, and dynamic.

Having established a set of design considerations, a comprehensive review of related work was conducted. For the three input measures this review covered measurement methods, feature extraction, and related research. With respect to output modalities, work on the minimal and dynamic visualisation of behaviour and affect was reviewed.

The second half of Chapter 4 documented the detailed design of the LuminUs: a device that was inspired and informed by the work reviewed earlier in the chapter, as well as the experiences and findings of the first study. In doing so, this contributed towards addressing the second aim of this thesis:

A2 Design and evaluate a sensor-based device for enhancing affective and behavioural interaction during collaborative music making.

The LuminUs uses eye-tracking headsets and worn accelerometers to provide musicians with visual feedback about the gaze or body motions of their co-performers. This feedback is presented visually using a strip of 16 coloured LEDs, which can be mounted on a flexible arm in front of the musician. Custom software was developed to collect and process the data; and to send lighting control messages to a small microcontroller inside the device. ECG sensors are also incorporated into the design of the LuminUs; enabling cardiac activity to be recorded as a latent measure (one that does not influence the visual feedback). This was included to facilitate further investigation of relationships between cardiac activity and musical decision making; which had been identified during the first study.

8.1.3 Study 2: Testing the LuminUs in the Lab (Chapter 5)

The second study was designed to test the LuminUs under controlled settings. The study aimed to i) investigate the effects of the LuminUs feedback upon collaborative music making; and ii) test for relationships between cardiac activity and musical decision making. The latter aim was justified in view of the potential for cardiac activity to be used to provide feedback in future versions of the LuminUs. For each aim, specific exploratory questions and hypotheses were tested. The study contributed towards addressing the second aim of this thesis:

A2 Design and evaluate a sensor-based device for enhancing affective and behavioural interaction during collaborative music making.

Fifteen pairs of musicians took part in the study, each consisting of a percussionist and a pianist. Each pair were asked to create an improvised accompaniment to a two minute animation under seven different LuminUs feedback conditions, with two attempts per condition. Gaze, cardiac, and motion measurements were collected from the musicians using eye-tracking headsets and wearable ECG sensors containing accelerometers. Self-report data were also collected, using questionnaires that were completed after each performance condition.

Findings relating to the effects of the LuminUs feedback showed that providing gaze feedback had significant influences upon the number of glances exchanged between participants; and the proportion of reciprocated glances, relative to the conditions where no feedback was provided. Motion feedback also had a positive influence on the number of glances exchanged. Neither gaze nor motion feedback had any significant influence upon the body motions of the performers, or self-reported aspects of the musical outcomes and creativity of the accompaniments.

A comprehensive analysis of relationships between cardiac activity and musical decision making was performed. This involved the use of two novel measures of musical decision making – *information content* and *change points* – which were automatically extracted from MIDI data. Both discrete and time series analyses were undertaken. The findings revealed only limited relationships between cardiac activity and musical decision making. In particular, the discrete analyses revealed that for pianists both the LF/HF ratio, and the number of heart rate (HR) extrema were significantly correlated with the mean information content. Visual analyses of the time series suggested that HR extrema coincided with animation cue points. However, cross-correlation analyses did not reveal any significant relationships between cardiac feature time series and decision making feature time series.

8.1.4 Study 3: The LuminUs in Practice (Chapter 6)

The final study in this thesis consisted of a longitudinal, qualitative study with four established musical duos. The main aims of the study were to investigate musicians' use of the LuminUs over a more sustained period, in more naturalistic settings; and to obtain qualitative findings that might help explain and contextualise the quantitative findings of the preceding studies. In this sense, the study contributed towards the second aim of this thesis (**A2**), whilst also addressing the third:

A3 Provide insights into the wider applications for affect and behaviour sensing technologies in the context of collaborative music making.

The duos attended four sessions where they rehearsed, composed, and performed with the LuminUs. Semi-structured interviews and video observations were used to collect qualitative data, which were subsequently transcribed, coded and analysed using thematic analysis. This resulted in the identification of five themes: i) expression and communication; ii) group attributes; iii) technology; iv) aspects of music; and v) context. These themes provided insights into the musicians' appropriation and use of the LuminUs; the interpretations of, and reactions to the two types of feedback (gaze and motion); and the overall impact of the device. Design considerations also emerged from these themes; leading to recommendations concerning the design of sensor-based devices to enhance affective and behavioural interaction during collaborative music making.

Across the five themes, findings were highlighted in relation to quantitative findings obtained in studies 1 and 2. Musicians' discussions surrounding interpretations of expressive motion highlighted that they associate body motion with the content of the music; and use it to gauge whether their co-performer is "into the music". This finding supported correlations between body motion and self-reported engagement and energy, observed in Study 1. When receiving motion feedback, musicians discussed using glances to compare the actual motions of their co-performer to the visual feedback provided by the LuminUs. This served as a potential explanation for the positive influence of motion feedback upon glances, observed in Study 2. Furthermore, musicians frequently discussed using gaze for signalling and cueing changes in the music. This supported quantitative associations between gaze and decision making, observed during Study 2.

8.2 Contributions

An in-depth discussion of the contributions made by this work is given below. As in Section 1.4, these are categorised in relation to i) **new methods proposed** and ii) **findings obtained**.

8.2.1 New Methods Proposed

The research in this thesis draws upon existing methods used in the fields of affective computing, social signal processing, psychophysiology, and musical interaction. In addition to combining and adapting these existing methods, this thesis also contributes to the aforementioned fields by developing, applying, and evaluating new methods for the purpose of understanding and enhancing affective and behavioural communication during collaborative interactions. Specific methods are discussed below:

Automatic tracking of human-human gaze interactions: Using open source eye tracking glasses and software, along with fiducial markers, a low-cost method was developed (Chapter 4) and evaluated (chapters 5 and 6) for automatically detecting gaze interactions between collaborating musicians. To the researcher's knowledge, this is the first time that eye-tracking has been adopted for this purpose. The method not only allowed the provision of real-time feedback to musicians about their co-performer's gaze, but also enabled the automatic collection of quantitative data on gaze interactions. Gaze plays an important role in the non-verbal communication of affective and social signals, and this method provides researchers with a tool for obtaining quantitative measures of gaze interactions in a range of different scenarios. For example, researchers in the field of computer supported cooperative work (CSCW) could use the gaze tracking method proposed in this thesis for investigating gaze interactions during collaborative problem solving.

Extracting quanititative measures of musical decision making: In Chapter 5, two novel methods were reported, applied, and evaluated for the automatic extraction of musical decision making features from MIDI recordings of improvised piano and drum performances. The first method used the Information Dynamics of Music (IDyOM) model (Pearce, 2005) to extract the information content (IC) - a representation of the unexpectedness of each musical the note. To the researcher's knowledge, the implementation of this method for the analysis of musical decision making in improvised music is a first. The second method involved the detection of musical change points using a combination of MIDI features, and was specifically developed for the purposes of this research. Both methods were evaluated in relation to the findings obtained during Study 2; indicating that the features provide descriptive and valuable information about musical decision making. When investigating collaborative musical interactions, it is important for researchers to take into account the musical contributions of the performers. This is especially challenging when the musicians are improvising, due to the fact that the musical contributions are not based upon a pre-defined score. The two methods reported in this thesis provide potentially valuable tools to enable researchers to extract quantitative indicators of the musical decisions made by musicians during improvised performances.

Providing visual feedback of affective and behavioural measures: By designing and testing a prototype device - the LuminUs - this thesis proposed and evaluated a method for providing musicians with real-time inter-personal feedback about the gaze and body motions of their co-performers. To the researcher's knowledge, this is the first time that sensor technologies have been used to provide such feedback to musicians during co-present performances. The method used a simple display, consisting of a strip of coloured LEDs, which was inspired by a review of existing methods for visually representing dynamic information.

The findings from studies 2 and 3 (chapters 5 and 6) indicated that, whilst the musicians did not find the display overly complex or distracting, its feedback was still capable of having a significant impact upon their interactions. Furthermore, the findings from Study 3 showed the importance of choosing suitable mappings between affective/behavioural signals and their associated visual representations. Consequently, the reported findings and experiences with the LuminUs provide an informative reference for researchers and designers involved in the development of devices for providing dynamic visual feedback to performing musicians.

8.2.2 Findings Obtained

In addition to the methodological contributions outlined above, this thesis also contributes towards advancing knowledge in multiple disciplines through the reporting of specific findings. Key findings are discussed below:

Sensing collaborating musicians: Findings from the first study in this thesis (Chapter 3) provide empirical evidence for relationships between sensor-derived features and self-reported, or performance-related measures of collaborative music making.

- Self-reported measures of creativity, engagement, and energy were found to be positively correlated with body motion.
- The number of glances exchanged between musicians were positively correlated with rhythmic synchrony.
- The average glance length was correlated with self-reported boredom.

These findings contribute towards future research by providing an informed indication of which sensors and features might be best suited for the automatic measurement and interpretation of affective and behavioural aspects of collaborative music making.

Musical decision making and cardiac activity: In Study 1 (Chapter 3), relationships between discrete heart rate and musical change events were evidenced. Further findings in Study 2 (Chapter 5) shed light upon relationships between cardiac activity and musical decision making, using both discrete and continuous time series analyses. These results contribute towards the wider literature on psychophysiology and sow the seeds for more in-depth future research on physiological indicators of decision making processes in improvising musicians.

Musicians' use of the LuminUs: Through the use of the LuminUs, quantitative evidence was obtained for the ways in which sensor-derived, visual feedback devices might have an impact upon musical collaborations. In particular, both gaze and motion feedback were shown to significantly influence the number of glances exchanged between musicians (Chapter 5). In Chapter 6 a longitudinal, qualitative study on the use of the LuminUs identified and evidenced five themes: expression and communication; group attributes; technology; aspects of music; and context. These findings contribute distinct recommendations and considerations for the design and development of affective and behavioural feedback devices for musicians.

Research and technology design recommendations: Based upon the experiences and findings throughout the course of this research, Chapter 7 provides a detailed discussion of recommendations and suggestions relating to the following topics:

- The selection of appropriate sensors for measuring behaviour and affect in collaborating musicians.
- The collection and analysis of dyadic data.
- Implications for the design of new affective and behavioural sensing technologies for collaborative music making.

The discussion of these topics contributes a concise and informative reference for researchers and technologists embarking upon work involving the use of sensors for measuring affective and behavioural aspects of collaborative music making. Additionally, it may also benefit those working within the wider context of the performing arts.

8.3 Limitations

This section briefly highlights limitations of the research undertaken in this thesis. Where possible, suggestions are provided for ways in which these limitations could be addressed in future work.

8.3.1 Study Groups

Collaborative music making can involve any number of musicians coming together to compose or perform music. The studies within this thesis were specifically limited to the investigation of pairs of musicians. This was justified by the need to constrain the breadth of the research, on account of the unexplored and complex nature of the research domain. Some aspects of collaborative musical interactions are likely to prevail regardless of group size. For example, musicians in Study 3 recounted a common use of gaze for cueing and signalling in groups of various sizes. However, it is not possible to confirm the extent to which specific findings in this thesis would be applicable to larger groups.

The range of instruments and styles of music played by the experimental dyads were also limited. Indeed, only five unique types of instrument were played during the studies: percussion, piano, electric guitar, bass guitar, and drum machine. Again, it is likely that certain aspects of collaborative musical interactions are common across a range of instruments and musical styles. However, the findings in Study 3 demonstrated the influence that certain musical situations, styles, and group attributes can have upon collaborative music making. These factors were particularly relevant to the ways in which the participants envisaged using the LuminUs. Finally, the participants investigated in this thesis were all experienced musicians. Therefore, the findings may not be applicable to interactions between amateur musicians.

To address the limitations highlighted in this section, it would be necessary to carry out further studies involving larger and more varied groups of musicians, playing a wider range of musical instruments and styles.

8.3.2 Statistical Analyses

The informativeness and validity of statistical analyses are determined by the ways in which data are collected, and by the selection and application of appropriate statistical techniques. A range of statistical methods were used for the analysis of quantitative data in the first two studies in this thesis. In Study 1, the participant sample size; ordering of the visibility condition; and validity of the post-performance questionnaire were highlighted as potential limitations that may have influenced the strength and validity of the statistical results. Consequently, further studies were recommended as a means of verifying the repeatability of these results.

An additional issue is the use of bivariate (one factor at-a-time) analyses. In the first study this was justified by the exploratory nature of the study and the large number of features that were analysed (see Section 3.5.2 for further discussion). During Study 2 the bivariate approach was predominantly used for investigating correlations between cardiac activity and musical decision making features. This was justified by the fact that little was known about the nature of any potential relationships; therefore, a bottom-up approach to the analyses was chosen. Despite these justifications, the use of bivariate analysis has a number of limitations, which are described in detail by Kent (2015). At a descriptive level, bivariate analysis precludes the identification and analysis of joint effects. For example, the findings from Study 1 showed that energy and engagement were both independently correlated with the mean body motion; however, bivariate analysis could not indicate the extent to which energy and engagement jointly related to body motion. At a relational level, bivariate analysis is unable to reveal conditional, confounding, or mediated relationships; which may exist when more than two variables are involved. For example, in Study 2 a positive relationship between the LuminUs motion feedback and the performers' body motions was identified. However, the motion feedback and performers' body motions were found to be confounded by the co-performers' body motions. In this case the confounding relationship was identified manually, due to an awareness of potential relationships between the variables. In future work, the use of multivariate analyses could be used in order to gain a more complex and thorough understanding of the relationships between groups of features.

One of the main limitations of statistical methods is that they are incapable of providing qualitative insights into data. This is particularly relevant to studies of human behaviour and psychology, where subjective and qualitative accounts of experience and action are an important consideration. In Study 2, quantitative findings evidenced the effects of the LuminUs feedback upon glancing behaviours; as well as relationships between the feedback and decision making features. However, it was only possible to speculate as to the causal pathways that led to these relationships, due to a lack of qualitative data. This was addressed to some extent by the qualitative findings obtained in Study 3.

Despite the issues discussed above, quantitative and statistically-based analysis methods are essential tools for extracting meaningful information from sensor data. The limitations of these methods can be addressed through the careful design of experiments; selection of appropriate tests; and supplementary use of qualitative research.

8.3.3 Sensor Technologies

The research in this thesis revolves around the use of sensor technologies, both for the collection of empirical data, and for the design of a real-time feedback device. Over the course of this research, various limitations associated with the sensor technologies were identified. In particular, these concern the practicalities of collecting measurements, and the reliability and informativeness of resulting data. These are discussed in more detail in Section 7.1. The rapid pace of technology development means that sensors used in the first study in this thesis are now out-dated, and have been replaced by more advanced technologies. Consequently, the justifications for not using certain sensors, due to issues such as unreliability or noise, may no longer be relevant, due to the existence of more capable devices. For example, in Study 1 the Emotiv EPOC headset was used to collect EEG data, but was retracted from further use due to the difficulties of setting up the headset, and the susceptibility of the readings to noise. However, Emotiv have since developed an EEG headset[1] that is specifically designed for everyday use, eliminating the need for extensive preparation, and using inertial sensors to reduce EEG signal noise (Emotiv, 2016).

8.3.4 External and Ecological Validity

In the context of human interaction research, external validity concerns the generalisability of research findings to other persons, groups and situations (Brewer and Crano, 2000). As discussed in Section 8.3.1, the studies undertaken in this thesis were constrained to dyadic interactions, involving a sub-set of musical instruments and styles. Furthermore, the first two studies used specific experimental tasks, which were performed in controlled settings. Consequently, the external validity of the findings obtained in these studies may be limited.

The ecological validity of a study concerns the degree to which the experimental conditions are representative of real-world conditions. In studies 1 and 2, attempts to obtain ecological validity were limited by the need to satisfy requirements for statistical validity. For example, stage lighting was used in both studies in an attempt to mimic real-world performance environments; but the experimental tasks, such as drumming with one hand, were not entirely representative of real-world musical interactions. This conflict is discussed further in sections 3.5.1 and 7.2.1.

8.4 Future Work

In this final section, a number of interesting areas for future work are briefly described. These present opportunities to advance, verify, explore and re-imagine the research reported in this thesis.

8.4.1 Taking the LuminUs into the Wild

The LuminUs is designed to be a device that can enhance collaborative music making activities, whilst also demonstrating potential applications for affect and behaviour

[1] `https://emotiv.com/insight.php`

sensing technologies. Within the scope of this thesis it was only possible to test the LuminUs in studies that were held inside university performance spaces. Participants in Study 3 suggested that the LuminUs could be beneficial in specific settings, such as an orchestral pit or recording studio. Consequently, it would be interesting to take the LuminUs into these environments and work with the musicians in order to explore and investigate specialist applications. This could also involve extending the focus from dyads to larger groups, such as orchestras. For example, each musician in an orchestra could be provided with a LuminUs, which informed them whenever the conductor was looking in their direction.

Taking the LuminUs into the wild would also provide an opportunity to collect and analyse data from naturalistic settings. This would help to address limitations relating to external and ecological validity, highlighted in the previous section. It would also present new challenges, such as dealing with data that was not subject to the same control that is afforded by lab-based studies.

8.4.2 Giving the LuminUs Intelligence

Volpe (2003) discusses three levels at which expressive input data can be mapped to expressive outputs: i) expressive direct mapping; ii) expressive high level mapping; and iii) expressive mapping monitoring (see Section 4.2.2 for a more detailed description). In its current state, the LuminUs provides a *direct mapping* of input measures to output visualisations (i.e. the data are processed using pre-defined computational operations that do not change state over time). This mapping was chosen in order to provide simple and meaningful feedback to the musicians. However, future work could look towards high level mappings (i.e. those that involve decision-making and reasoning processes). This could be facilitated by developments in the fields of machine learning and artificial intelligence, which offer powerful methods for the automatic interpretation and recognition of patterns in data. These methods have already been adopted in the fields of affective computing and social signal processing as a means of automatically recognising affective and social phenomena (Sariyanidi et al., 2015; Vinciarelli et al., 2009; Zeng et al., 2009). Considering the motion and gaze data available to the LuminUs, there is great potential for high level mapping strategies to be developed; expanding the capabilities of the device, and giving it 'intelligence'. This could involve combining data from multiple sensors. For example, a musician's motion data could be used to provide an indication of whether they were currently playing or not; allowing the LuminUs to recognise whether a glance from their co-performer resulted in them changing their playing. Subsequently, the LuminUs could use this information to learn about which combinations of gaze and motion were associated with functional cueing glances, compared to more general expressive glances.

8.4.3 Utilising Wearable and Everyday Sensors

The sensors used in this thesis would not be classified as everyday, consumer technologies. By undertaking research with widely used technologies, such as smartphones, the research outcomes would be more accessible to the general population. For example, it would be feasible to develop a LuminUs app for smartphones, which enabled groups of musicians to wirelessly link their phones and use the screens to display motion feedback, derived from the accelerometers inside the phones. Furthermore, this would facilitate the collection of data from real-world collaborative music making activities; since participants could simply use their own devices, without any need for supervision from the researchers.

An exciting area of consumer technology development is wearable technology. 'Wearables' – as they are commonly referred to – often include similar sensors to those used in this thesis. For example, the Apple Watch can measure the wearer's heart rate from their wrist, and incorporates an accelerometer (Limer, 2014). It can also provide haptic feedback through vibrations. Consequently, the Apple Watch could easily collect measures from performing musicians and provide them with discrete vibro-tactile feedback. Such devices provide promising opportunities for researchers to use sensor data for undertaking investigations and developing applications involving affective and behavioural signals.

8.4.4 Designing Affective and Intelligent Musical Instruments

Devices like the LuminUs have the potential to provide collaborating musicians with an enhanced awareness of affective and behavioural signals. But what if these signals also influenced the functioning of a collaborative musical instrument? For example, an electronic instrument could use motion data to infer levels of engagement and, rather than informing the musician that their collaborator was no longer engaged, it could subtly suggest or make changes to the music. Equipping musical interfaces with the ability to sense behavioural and physiological phenomena is not an entirely new idea. Indeed, new interfaces for musical expression often incorporate sensor technologies as a means of generating sound (Medeiros and Wanderley, 2014). However, the utilisation of these sensors for affect and social signal sensing purposes has not been well explored. Consequently, in many new interfaces for musical expression the hardware capabilities for data collection may already exist; meaning that it would only be necessary to modify or develop software in order to process these data. Where more widely used interfaces are concerned, this could facilitate the collection of large data sets from people using these devices in the wild. Furthermore, musicians who choose to adopt new musical interfaces may be more receptive to new technologies in general; meaning that they could be more likely to adopt and benefit from devices for supporting affective and inter-personal aspects of collaborative music making.

8.4.5 Assisting Musicians with Disabilities

The LuminUs provides performing musicians with feedback about the gaze or body motions of their co-performers. For musicians lacking in mobility, devices like the LuminUs could provide a particularly useful means of facilitating non-verbal expression and communication. For example, musicians in Study 3 discussed using movement as a means of identifying whether a co-performer was 'into the music'. Consequently, where musicians are unable to express visible body movements, smaller movements could be sensed and represented through visual feedback. This feedback could be presented to the audience, as well as other performers. Study 3 also highlighted the importance of attracting the attention of other musicians for signalling and cueing purposes. The use of gaze detection provides musicians with a simple and direct way of getting their co-performer's attention without the need for elaborate gestures. Furthermore, people with significant physical impairments are often already equipped with eye-tracking technologies as a means of communication. These technologies have been utilised for music making purposes (Hornof, 2014). However, there are no known cases where eye-tracking has been used for enhancing interaction between musicians with disabilities.

8.5 Closing Remarks

Instrumenting the Musician – the title of this thesis – denotes the use of instruments, not as tools for musical creation, but as tools for measuring, understanding, and enhancing musical interaction. Over the course of this research sensor technologies have been investigated and applied to the measurement and enhancement of affective and behavioural interactions between collaborating musicians. In a year referred to as the "year when virtual reality goes from virtual to reality" (Cellan-Jones, 2016), it is hoped that this work will serve to highlight the ways in which technology can also contribute to the understanding and enrichment of the behaviourally and emotionally diverse reality of co-present, human-human interaction.

Bibliography

Anne Adams, Peter Lunt, and Paul Cairns. A qualitative approach to HCI research. In Paul Cairns and Anna L. Cox, editors, *Research methods for human-computer interaction*, chapter 7, pages 138–157. Cambridge University Press, Cambridge, UK, 2008. ISBN 9780521690317. URL `http://www.cambridge.org/gb/academic/subjects/psychology/applied-psychology/research-methods-human-computer-interaction`.

Reginald B Adams and Robert E Kleck. Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion (Washington, D.C.)*, 5(1): 3–11, March 2005. ISSN 1528-3542. doi: 10.1037/1528-3542.5.1.3. URL `http://www.ncbi.nlm.nih.gov/pubmed/15755215`.

Teresa M. Amabile. Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43(5):997–1013, 1982. ISSN 0022-3514. doi: 10.1037//0022-3514.43.5.997. URL `http://content.apa.org/journals/psp/43/5/997`.

Teresa M. Amabile, Sigal G. Barsade, Jennifer S. Mueller, and Barry M. Staw. Affect and Creativity at Work. *Administrative Science Quarterly*, 50(3):367–403, September 2005. ISSN 0001-8392. doi: 10.2189/asqu.2005.50.3.367. URL `http://asq.sagepub.com/content/50/3/367.short`.

Yair Amichai-Hamburger and Zack Hayat. The impact of the Internet on the social lives of users: A representative sample from 13 countries. *Computers in Human Behavior*, 27(1):585–589, January 2011. ISSN 07475632. doi: 10.1016/j.chb.2010.10.009. URL `http://linkinghub.elsevier.com/retrieve/pii/S0747563210003067`.

R J Anderson. Representations and Requirements: The Value of Ethnography in System Design. *Human-Computer Interaction*, 9(2):151–182, June 1994. ISSN 0737-0024. doi: 10.1207/s15327051hci0902{_}1.

Bradley M. Appelhans and Linda J. Luecken. Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, 10(3):229–240, 2006. ISSN 1939-1552. doi: 10.1037/1089-2680.10.3.229.

Michael Argyle. *The psychology of interpersonal behaviour*. Penguin Books, 1978.

Michael Argyle and Jean Ann Graham. The central Europe experiment: Looking at persons and looking at objects. *Environmental Psychology and Nonverbal Behavior*, 1(1):6–16, 1976. ISSN 0361-3496. doi: 10.1007/BF01115461.

Michael Argyle, Roger Ingham, Florisse Alkema, and Margaret McCallin. The Different Functions of Gaze. *Semiotica*, 7(1), 1973. ISSN 0037-1998. doi: 10.1515/semi.1973.7.1.19. URL `http://www.degruyter.com/view/j/semi.1973.7.issue-1/semi.1973.7.1.19/semi.1973.7.1.19.xml`.

R.H. Baayen, D.J. Davidson, and D.M. Bates. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4):390–412, November 2008. ISSN 0749596X. doi: 10.1016/j.jml.2007.12.005. URL `http://linkinghub.elsevier.com/retrieve/pii/S0749596X07001398`.

Claudio Babiloni, Paola Buffo, Fabrizio Vecchio, Nicola Marzano, Claudio Del Percio, Danilo Spada, Simone Rossi, Ivo Bruni, Paolo M Rossini, and Daniela Perani. Brains "in concert": frontal oscillatory alpha rhythms and empathy in professional musicians. *NeuroImage*, 60(1):105–16, March 2012. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2011.12.008. URL `http://www.ncbi.nlm.nih.gov/pubmed/22186679`.

S.G. Barsade. The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, (December):644–675, 2002. doi: 10.2307/3094912.

a Bechara, H Damasio, D Tranel, and a R Damasio. The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends in cognitive sciences*, 9(4):159–62; discussion 162–4, April 2005. ISSN 1364-6613. doi: 10.1016/j.tics.2005.02.002. URL `http://www.ncbi.nlm.nih.gov/pubmed/15808493`.

Roman Bednarik, Andrey Shipilov, and Sami Pietinen. Bidirectional gaze in remote computer mediated collaboration. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work - CSCW '11*, page 597, New York, New York, USA, 2011. ACM Press. ISBN 9781450305563. doi: 10.1145/1958824.1958923. URL `http://portal.acm.org/citation.cfm?doid=1958824.1958923`.

Daniel Belenky, Michael Ringenberg, and Jennifer Olsen. Using Dual Eye-Tracking to Evaluate Students' Collaboration with an Intelligent Tutoring System for Elementary-Level Fractions. In *Proceedings of Cognitive Science (CogSci 2014)*, pages 176–181, 2014. URL `https://mindmodeling.org/cogsci2014/papers/041/paper041.pdf`.

Erwin B Bellers, Johan G W M Janssen, and Maurice Penners. 44.4: Motion Compensated Frame Rate Conversion for Motion Blur Reduction. *SID Symposium Digest of Technical Papers*, 38(1):1454–1457, May 2007. ISSN 0097966X. doi: 10.1889/1.2785589.

Mathias Benedek and Christian Kaernbach. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1):80–91, June 2010. ISSN 1872-678X. doi: 10.1016/j.jneumeth.2010.04.028. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2892750&tool=pmcentrez&rendertype=abstract`.

Ben Bengler and Nick Bryan-Kinns. Designing collaborative musical experiences for broad audiences. *Proceedings of the 9th ACM Conference on Creativity & Cognition - C&C '13*, page 234, 2013. doi: 10.1145/2466627.2466633. URL `http://dl.acm.org/citation.cfm?doid=2466627.2466633`.

M.S. Benning, a. Kapur, B. Till, G. Tzanetakis, and P. Driessen. A Comparative Study on Wearable Sensors for Signal Processing on the North Indian Tabla. *2007 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, (January):1–4, 2007.

L Bernardi, C Porta, and P Sleight. Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence. *Heart (British Cardiac Society)*, 92(4):445–52, April 2006. ISSN 1468-201X. doi: 10.1136/hrt.2005.064600. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1860846&tool=pmcentrez&rendertype=abstract`.

G G Berntson, J T Bigger, D L Eckberg, P Grossman, P G Kaufmann, M Malik, H N Nagaraja, S W Porges, J P Saul, P H Stone, and M W van der Molen. Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, 34(6):623–48, November 1997. ISSN 0048-5772. URL `http://www.ncbi.nlm.nih.gov/pubmed/9401419`.

Gary G. Berntson, Karen S. Quigley, and Dave Lozano. Cardiovascular Psychophysiology. In John T. Cacioppo, Louis G. Tassinary, and Gary Berntson, editors, *Handbook of Psychophysiology*, pages 182–210. Cambridge University Press, Cambridge, 2007. doi: 10.1017/CBO9780511546396.008. URL `http://ebooks.cambridge.org/ref/id/CBO9780511546396A016`.

Nadia Bianchi-Berthouze. Understanding the Role of Body Movement in Player Engagement. *HumanComputer Interaction*, 23(1):40–75, 2012. doi: 10.1080/07370024.2012.688468.

Nadia Bianchi-Berthouze and Andrea Kleinsmith. Automatic Recognition of Affective Body Expressions. In Rafael Calvo, Sidney D'Mello, Jonathan Gratch, and Arvid Kappas, editors, *The Oxford Handbook of Affective Computing*. Oxford University Press, January 2015. doi: 10.1093/oxfordhb/9780199942237.013.025.

Michele Biasutti. Group music composing strategies: A case study within a rock band. *British Journal of Music Education*, (August):1–15, 2012. ISSN 0265-0517. doi: 10.1017/S0265051712000289.

Michele Biasutti and Luigi Frezza. Dimensions of Music Improvisation. *Creativity Research Journal*, 21(2-3):232–242, May 2009. ISSN 1040-0419. doi: 10.1080/10400410902861240.

Kenneth C Bilchick and Ronald D Berger. Heart rate variability. *Journal of cardiovascular electrophysiology*, 17(6):691–4, June 2006. ISSN 1045-3873. doi: 10.1111/j.1540-8167.2006.00501.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/16836727`.

Ann Blandford. Semi-Structured Qualitative Studies. In Mads Soegaard and Rikke Friis Dam, editors, *The Encyclopedia of Human-Computer Interaction*. The Interaction Design Foundation, Aarhus, Denmark, 2013.

J Blascovich and WB Mendes. Social psychophysiology and embodiment. *Handbook of Social Psychology*, pages 194–227, 2010. doi: 10.1002/9780470561119.socpsy001006/full.

K S Bowers and K R Keeling. Heart-rate variability in creative functioning. *Psychological reports*, 29(1):160–2, August 1971. ISSN 0033-2941. doi: 10.2466/pr0.1971.29.1.160. URL `http://www.ncbi.nlm.nih.gov/pubmed/4937323`.

Radu Gabriel Bozomitu, Alexandru Pasarica, Vlad Cehan, Cristian Rotariu, and Constantin Barabasa. Pupil centre coordinates detection using the circular Hough transform technique. In *2015 38th International Spring Seminar on Electronics Technology (ISSE)*, number 1, pages 462–465. IEEE, May 2015. ISBN 978-1-4799-8860-0. doi: 10.1109/ISSE.2015.7248041. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7248041`.

Margaret M. Bradley and Peter J. Lang. Emotion and Motivation. In John T. Cacioppo, Louis G. Tassinary, and Gary Berntson, editors, *Handbook of Psychophysiology*, pages 581–607. Cambridge University Press, Cambridge, 2007. doi: 10.1017/CBO9780511546396.025. URL `http://ebooks.cambridge.org/ref/id/CBO9780511546396A034`.

V. Braun and V. Clarke. Using Thematic Analysis in Psychology. *Qualitative research in psychology*, 3(2):77–101, 2006. ISSN 1478-0887. doi: 10.1191/1478088706qp063oa.

Marilynn B Brewer and William D. Crano. Research Design and Issues of Validity. In Harry T. Reis and Charles M. Judd, editors, *Handbook of Research Methods in Social and Personality Psychology*, pages 11–26. Cambridge University Press, New York, 2000. ISBN 0521559030. doi: 10.1017/CBO9780511996481.005. URL `http://ebooks.cambridge.org/ref/id/CBO9780511996481A011`.

El Broek, Jh Janssen, Jhdm Westerink, and Ja Healey. Prerequisites for affective signal processing (ASP). In *International Conference on Bio-Inspired Systems and Signal Processing, Biosignals 2009*, number Box 1, 2009. ISBN 9789896740184. URL `http://doc.utwente.nl/73169/`.

Nick Bryan-Kinns. Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies*, 71(1):76–90, January 2013. ISSN 10715819. doi: 10.1016/j.ijhcs.2012.02.004. URL `http://linkinghub.elsevier.com/retrieve/pii/S1071581912000328`.

Nick Bryan-Kinns and Fraser Hamilton. Identifying mutual engagement. *Behaviour & Information Technology*, 31(2):101–125, February 2012. ISSN 0144-929X. doi: 10.1080/01449290903377103.

TH Budzynski, HK Budzynski, JR Evans, and A Abarbanel. *Introduction to Quantitative EEG and Neurofeedback: Advanced Theory and Applications*. Academic Press, 2nd edition, 2008. ISBN 9780123745347.

John T. Cacioppo, Louis G. Tassinary, and Gary Berntson, editors. *Handbook of Psychophysiology*. Cambridge University Press, Cambridge, 2007a. ISBN 9780511546396.

doi: 10.1017/CBO9780511546396. URL http://ebooks.cambridge.org/ref/id/CBO9780511546396.

John T. Cacioppo, Louis G. Tassinary, and Gary G. Berntson. Psychophysiological Science: Interdisciplinary Approaches to Classic Questions About the Mind. In John T. Cacioppo, Louis G. Tassinary, and Gary Berntson, editors, *Handbook of Psychophysiology*, pages 1–16. Cambridge University Press, Cambridge, 2007b. doi: 10.1017/CBO9780511546396.001. URL http://ebooks.cambridge.org/ref/id/CBO9780511546396A008.

Antonio Camurri, Ingrid Lagerlöf, and Gualtiero Volpe. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, 59(1-2):213–225, July 2003. ISSN 10715819. doi: 10.1016/S1071-5819(03)00050-8. URL http://linkinghub.elsevier.com/retrieve/pii/S1071581903000508.

Antonio Camurri, Barbara Mazzarino, and Gualtiero Volpe. Analysis of expressive gesture: The eyesweb expressive gesture processing library. *Gesture-based communication in ...*, pages 460–467, 2004. doi: 10.1007/978-3-540-24598-8_42. URL http://www.springerlink.com/index/RFT1L2J1UM7W2H86.pdf.

Jennifer Carlile and Bjorn Hartmann. OROBORO : A Collaborative Controller with Interpersonal Haptic Feedback. In *New Interfaces for Musical Expression (NIME05)*, pages 250–251, 2005.

M.A. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney. Content-Based Music Information Retrieval: Current Directions and Future Challenges. *Proceedings of the IEEE*, 96(4):668–696, 2008. ISSN 0018-9219. doi: 10.1109/JPROC.2008.916370. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4472077.

Ginevra Castellano, Santiago D. Villalba, and Antonio Camurri. Recognising Human Emotions from Body Movement and Gesture Dynamics. In *Affective Computing and Intelligent Interaction (ACII07)*, pages 71–82, Lisbon, Portugal, 2007. Springer-Verlag. ISBN 2002507422. doi: 10.1007/978-3-540-74889-2_7. URL http://www.springerlink.com/index/aj246lvw0635g238.pdf.

Ginevra Castellano, Marcello Mortillaro, Antonio Camurri, Gualtiero Volpe, and Klaus Scherer. Automated Analysis of Body Movement in Emotionally Expressive Piano

Performances. *Music Perception: An Interdisciplinary Journal*, 26(2):103–119, December 2008. ISSN 07307829. doi: 10.1525/mp.2008.26.2.103.

Rory Cellan-Jones. 2016: the year when VR goes from virtual to reality. Online, January 2016. URL http://www.bbc.co.uk/news/technology-35205783.

D Cernea, A Ebert, and A Kerren. Visualizing Group Affective Tone in Collaborative Scenarios. In N. Elmqvist, M. Hlawitschka, and J. Kennedy, editors, *Eurographics Conference on Visualization (EuroVis) (2014)*, 2014. URL http://homepage.lnu.se/staff/akemsi/pubs/kerren-eurovis14.pdf.

Daniel Cernea, Christopher Weber, Achim Ebert, and Andreas Kerren. Emotion scents: a method of representing user emotions on GUI widgets. In Pak Chung Wong, David L. Kao, Ming C. Hao, Chaomei Chen, and Christopher G. Healey, editors, *Proc. SPIE 8654, Visualization and Data Analysis 2013*, volume 8654, February 2013. doi: 10.1117/12.2001261. URL http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1568718.

B. Chamadiya, K. Mankodiya, M. Wagner, and U. G. Hofmann. Textile-based, contactless ECG monitoring for non-ICU clinical settings. *Journal of Ambient Intelligence and Humanized Computing*, 4(6):791–800, July 2012. ISSN 1868-5137. doi: 10.1007/s12652-012-0153-8.

Guillaume Chanel, Mireille Bétrancourt, Thierry Pun, Donato Cereghetti, and Gaëlle Molinari. Assessment of Computer-Supported Collaborative Processes using Interpersonal Physiological and Eye-Movement Coupling. In *Affective Computing and Intelligent Interaction (ACII13)*, Geneva, Switzerland, 2013. URL http://www.affective-sciences.org/system/files/biblio/ACII_2013_20130710_Correction_Final_3.pdf.

Chuan-Yu Chang, Jeng-Shiun Tsai, Chi-Jane Wang, and Pau-Choo Chung. Emotion recognition with consideration of facial expression and physiological signals. In *2009 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 278–283. IEEE, March 2009. ISBN 978-1-4244-2756-7. doi: 10.1109/CIBCB.2009.4925739. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4925739.

Maher Chaouachi and Claude Frasson. Exploring the relationship between learner EEG mental engagement and affect. In Vincent Aleven, Judy Kay, and Jack Mostow, editors, *Intelligent Tutoring Systems*, pages 291–293. Springer Berlin Heidelberg, 2010. doi: 10.1007/978-3-642-13437-1_48.

Kathy Charmaz. *Constructing Grounded Theory: A Practical Guide Through Qualitative Research*. SAGE Publications, 2006. ISBN 9780761973522.

Tanya L Chartrand and Jessica L Lakin. The Antecedents and Consequences of Human Behavioral Mimicry. *Annual Review of Psychology*, (September 2012), September 2012. ISSN 1545-2085. doi: 10.1146/annurev-psych-113011-143754. URL `http://www.ncbi.nlm.nih.gov/pubmed/23020640`.

Mauro Cherubini, Marc-Antoine Nüssli, and Pierre Dillenbourg. Deixis and gaze in collaborative work at a distance (over a shared map): a computational model to detect misunderstandings. In *Proceedings of the 2008 symposium on Eye tracking research & applications - ETRA '08*, volume 1, page 173, New York, New York, USA, 2008. ACM Press. ISBN 9781595939821. doi: 10.1145/1344471.1344515. URL `http://portal.acm.org/citation.cfm?doid=1344471.1344515`.

H.A. Chinn, D.K. Gannett, and R.M. Morris. A New Standard Volume Indicator and Reference Level. *Proceedings of the IRE*, 28(1):1–17, January 1940. ISSN 0096-8390. doi: 10.1109/JRPROC.1940.228815. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1687016`.

Ross a Clark, Yong-Hao Pua, Karine Fortin, Callan Ritchie, Kate E Webster, Linda Denehy, and Adam L Bryant. Validity of the Microsoft Kinect for assessment of postural control. *Gait & posture*, 36(3):372–7, July 2012. ISSN 1879-2219. doi: 10.1016/j.gaitpost.2012.03.033. URL `http://www.ncbi.nlm.nih.gov/pubmed/22633015`.

Tanya J Clarke, Mark F Bradshaw, David T Field, Sarah E Hampson, and David Rose. The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception*, 34(10):1171–1180, 2005. ISSN 0301-0066. doi: 10.1068/p5203. URL `http://www.perceptionweb.com/abstract.cgi?id=p5203`.

G.D. Clifford, P.E. McSharry, and L. Tarassenko. Characterizing artefact in the normal human 24-hour RR time series to aid identification and artificial replication of circadian variations in human beat to beat heart rate using a simple threshold. *Computers in Cardiology*, 2002(38):129–132, 2002. ISSN 0276-6547. doi: 10.1109/CIC.2002.1166724.

R Cowie, H Gunes, G Mckeown, L Vaclavu-Schneider, J Armstrong, and E Douglas-Cowie. The emotional and communicative significance of head nods and shakes in a naturalistic database. *International Workshop on Emotion*, pages 42–46, 2010.

Hewitt D Crane and Carroll M Steele. Generation-V dual-Purkinje-image eyetracker. *Applied Optics*, 24(4):527, February 1985. ISSN 0003-6935. doi: 10.1364/AO.24.000527. URL `http://www.ncbi.nlm.nih.gov/pubmed/18216982`.

Eveline a Crone, Riek J M Somsen, Bert Van Beek, and Maurits W Van Der Molen. Heart rate and skin conductance analysis of antecendents and consequences of decision making. *Psychophysiology*, 41(4):531–40, July 2004. ISSN 0048-5772. doi: 10.1111/j.1469-8986.2004.00197.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/15189476`.

Jonathan D Cryer and Kung-Sik Chan. *Time Series Analysis*. Number January in Springer Texts in Statistics. Springer New York, New York, NY, 2008. ISBN 978-0-387-75958-6. doi: 10.1007/978-0-387-75959-3.

Mihaly Csikszentmihalyi and Kevin Rathunde. *The measurement of flow in everyday life: Toward a theory of emergent motivation.*, volume 40, pages 57–97. University of Nebraska Press, Lincoln, NE, US, 1993.

Gabriela Csurka, Sandra Skaff, Luca Marchesotti, and Craig Saunders. Learning moods and emotions from color combinations. In *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing - ICVGIP '10*, pages 298–305, New York, New York, USA, 2010. ACM Press. ISBN 9781450300605. doi: 10.1145/1924559.1924599. URL `http://portal.acm.org/citation.cfm?doid=1924559.1924599`.

Fred Cummins. Gaze and blinking in dyadic conversation: A study in coordinated behaviour among individuals. *Language and Cognitive Processes*, 27(10):1525–1549, December 2012. ISSN 0169-0965. doi: 10.1080/01690965.2011.615220.

Jared R Curhan and Alex Pentland. Thin slices of negotiation: predicting outcomes from conversational dynamics within the first 5 minutes. *The Journal of applied psychology*, 92(3):802–11, 2007. ISSN 0021-9010. doi: 10.1037/0021-9010.92.3.802. URL `http://www.ncbi.nlm.nih.gov/pubmed/17484559`.

N. Dael, N. BianchiBerthouze, A. Kleinsmith, and C. Mohr. Measuring body movement: Current and future directions in proxemics and kinesics. *APA Handbook of Nonverbal Communication*, 2016.

Sofia Dahl and Anders Friberg. Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception: An Interdisciplinary Journal*, 24(5):433–454, June 2007. ISSN 07307829. doi: 10.1525/mp.2007.24.5.433.

Sofia Dahl, Frédéric Bevilacqua, Roberto Bresin, Martin Clayton, Laura Leante, Isabella Poggi, and Nicolas Rasamimanana. *Gestures in performance*. Routledge, 2010. ISBN 978-0-415-99886-4.

Mat Dalgleish and Steve Spencer. Postrum : Developing Good Posture in Trumpet Players Through Directional Haptic Feedback. *Proceedings of the 9th Conference on Interdisciplinary Musicology CIM14*, (October):38, 2015.

A R Damasio. The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 351(1346):1413–20, October 1996. ISSN 0962-8436. doi: 10.1098/rstb.1996.0125. URL `http://rstb.royalsocietypublishing.org/content/351/1346/1413.short`.

J. W. Davidson and J. M. M. Good. Social and Musical Co-Ordination between Members of a String Quartet: An Exploratory Study. *Psychology of Music*, 30(2):186–201, October 2002. ISSN 0305-7356. doi: 10.1177/0305735602302005.

Carsten K W De Dreu, Matthijs Baas, and Bernard a Nijstad. Hedonic tone and activation level in the mood-creativity link: toward a dual pathway to creativity model. *Journal of Personality and Social Psychology*, 94(5):739–56, May 2008. ISSN 0022-3514. doi: 10.1037/0022-3514.94.5.739. URL `http://www.ncbi.nlm.nih.gov/pubmed/18444736`.

Orjan de Manzano, Töres Theorell, László Harmat, and Fredrik Ullén. The psychophysiology of flow during piano playing. *Emotion (Washington, D.C.)*, 10 (3):301–11, June 2010. ISSN 1931-1516. doi: 10.1037/a0018432. URL `http://www.ncbi.nlm.nih.gov/pubmed/20515220`.

R. T. Dean and F. Bailes. Using time series analysis to evaluate skin conductance during movement in piano improvisation. *Psychology of Music*, September 2013. ISSN 0305-7356. doi: 10.1177/0305735613489917.

Roger T. Dean and William T. M. Dunsmuir. Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, June 2015. ISSN 1554-3528. doi: 10.3758/s13428-015-0611-2.

Gina Deininger and Gareth Loudon. Correlation between coherent heart rate variability and divergent thinking. In *Proceedings of the 8th ACM conference on Creativity and cognition - C&C '11*, pages 329–330, New York, New York, USA, 2011. ACM Press. ISBN 9781450308205. doi: 10.1145/2069618.2069681. URL `http://dl.acm.org/citation.cfm?doid=2069618.2069681`.

Arnaud Delorme and Scott Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, March 2004. ISSN 0165-0270. doi: 10.1016/j.jneumeth.2003.10.009. URL `http://www.ncbi.nlm.nih.gov/pubmed/15102499`.

Benjamin H. Detenber, Robert F. Simons, and Jason E. Reiss. The Emotional Significance of Color in Television Presentations. *Media Psychology*, 2(4):331–355, November 2000. ISSN 1521-3269. doi: 10.1207/S1532785XMEP0204_02.

Arne Dietrich and Riam Kanso. A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin*, 136(5):822–48, October 2010. ISSN 1939-1455. doi: 10.1037/a0019749. URL `http://www.ncbi.nlm.nih.gov/pubmed/20804237`.

Simon Dixon. Automatic Extraction of Tempo and Beat From Expressive Performances. *Journal of New Music Research*, 30(1):39–58, March 2001. ISSN 0929-8215. doi: 10.1076/jnmr.30.1.39.7119.

J. Stephen Downie. Music Information Retrieval. *Annual Review of Information Science and Technology*, 37(Chapter 7):295–340, 2003. ISSN 0093402X. doi: 10.1002/aris.1440370108.

Michelle Drouin and Claire Davis. R u txting? Is the Use of Text Speak Hurting Your Literacy? *Journal of Literacy Research*, 41(1):46–67, January 2009. ISSN 1086-296X. doi: 10.1080/10862960802695131. URL `http://jlr.sagepub.com/content/41/1/46.abstract`.

Starkey Duncan. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292, 1972. ISSN 0022-3514. doi: 10.1037/h0033031. URL `http://content.apa.org/journals/psp/23/2/283`.

Barnaby D Dunn, Tim Dalgleish, and Andrew D Lawrence. The somatic marker hypothesis: a critical evaluation. *Neuroscience and biobehavioral reviews*, 30(2):239–71, January 2006. ISSN 0149-7634. doi: 10.1016/j.neubiorev.2005.07.001. URL `http://www.ncbi.nlm.nih.gov/pubmed/16197997`.

Matthieu Duvinage, Thierry Castermans, Mathieu Petieau, Thomas Hoellinger, Guy Cheron, and Thierry Dutoit. Performance of the Emotiv Epoc headset for P300-based applications. *Biomedical Engineering Online*, 12:56, 2013. ISSN 1475-925X. doi: 10.1186/1475-925X-12-56. URL `http://www.ncbi.nlm.nih.gov/pubmed/23800158`.

Luke Edwards. Ford S-MAX Concept will measure your heart rate from the driver's seat. Online, September 2013. URL `http://www.pocket-lint.com/news/123510`. (Accessed 25th September 2013).

Hauke Egermann, Marcus T. Pearce, Geraint a. Wiggins, and Stephen McAdams. Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective & Behavioral Neuroscience*, 13 (3):533–53, September 2013. ISSN 1531-135X. doi: 10.3758/s13415-013-0161-y.

Emotiv. Emotiv Insight, 2016. URL `https://emotiv.com/insight.php`.

Stephen H. Fairclough. Fundamentals of physiological computing. *Interacting with Computers*, 21(1-2):133–145, 2009. ISSN 09535438. doi: 10.1016/j.intcom.2008.10.011.

Robin Fencott and Nick Bryan-Kinns. Hey Man , you re invading my Personal Space! Privacy and Awareness in Collaborative Music. In *New Interfaces for Musical Expression (NIME10)*, pages 198–203, 2010.

Uwe Flick. *An Introduction to Qualitative Research*, volume 4th. 2009. ISBN 9781847873248. URL `http://books.google.cz/books?id=sFv1oWX2DoEC`.

Wolfgang Fuhl, Thomas Kübler, Katrin Sippel, Wolfgang Rosenstiel, and Enkelejda Kasneci. ExCuSe: Robust Pupil Detection in Real-World Scenarios. In George Azzopardi and Nicolai Petkov, editors, *Computer Analysis of Images and Patterns*, pages 39–51. Springer International Publishing, 2015. ISBN 978-3-319-23191-4. doi: 10.1007/978-3-319-23192-1{_}4.

Liane Gabora. Cognitive mechanisms underlying the creative process. In *Proceedings of the fourth conference on Creativity & cognition - C&C '02*, pages 126–133, New York, New York, USA, 2002. ACM Press. ISBN 1581134657. doi: 10.1145/581710.581730. URL `http://portal.acm.org/citation.cfm?doid=581710.581730`.

Andrea Gaggioli, Pietro Cipresso, Silvia Serino, and Giuseppe Riva. Psychophysiological correlates of flow during daily activities. *Studies in health technology* *and informatics*, 191(10):65–9, January 2013. ISSN 0926-9630. doi: 10.3233/978-1-61499-282-0-65. URL http://www.ncbi.nlm.nih.gov/pubmed/23792845.

Philip Galanter. Computational Aesthetic Evaluation: Past and Future. In *Computers and Creativity*, pages 270–307. 2013. doi: 10.1007/978-3-642-31727-9{_}10.

A Gale, E Kingsley, S Brookes, and D Smith. Cortical arousal and social intimacy in the human female under different conditions of eye contact. *Behavioural processes*, 3 (3):271–5, October 1978. ISSN 0376-6357. doi: 10.1016/0376-6357(78)90019-0. URL http://www.ncbi.nlm.nih.gov/pubmed/24924822.

Brook Galna, Gillian Barry, Dan Jackson, Dadirayi Mhiripiri, Patrick Olivier, and Lynn Rochester. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait & posture*, 39(4):1062–8, April 2014. ISSN 1879-2219. doi: 10.1016/j.gaitpost.2014.01.008. URL http://www.ncbi.nlm.nih.gov/pubmed/24560691.

Nicholas D Giardino, Paul M Lehrer, and Robert Edelberg. Comparison of finger plethysmograph to ECG in the measurement of heart rate variability. *Psychophysiology*, 39(2):246–53, March 2002. ISSN 0048-5772. doi: 10.1017/S0048577202990049. URL http://www.ncbi.nlm.nih.gov/pubmed/12212675.

Christian Glahn, Marcus Specht, and Rob Koper. Visualisation of Interaction Footprints for Engagement in Online Communities. *Educational Technology & Society*, 12(3):44–57, 2009.

Peter A. Gloor, Daniel Oster, and Kai Fischbach. JazzFlowAnalyzing Group Flow Among Jazz Musicians Through Honest Signals. *KI - Künstliche Intelligenz*, 27(1): 37–43, 2013. ISSN 0933-1875. doi: 10.1007/s13218-012-0230-3.

Donald Glowinski, Nele Dael, Antonio Camurri, Gualtiero Volpe, Marcello Mortillaro, and Klaus Scherer. Toward a minimal representation of affective gestures. *IEEE Transactions on Affective Computing*, 2(2):106–118, 2011. ISSN 19493045. doi: 10.1109/T-AFFC.2011.7.

Donald Glowinski, Giorgio Gnecco, Stefano Piana, and Antonio Camurri. Expressive Non-Verbal Interaction in String Quartet. In *Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII'13)*, pages 233–238, Geneva, Switzerland, 2013. IEEE Computer Society. ISBN 9780769550480. doi: 10.1109/ACII.2013.45.

Werner Goebl and Caroline Palmer. Synchronization of Timing and Motion Among Performing Musicians. *Music Perception*, 26(5):427–438, 2009. ISSN 0730-7829. doi: 10.1525/mp.2009.26.5.427.

Werner Goebl, Simon Dixon, and Emery Schubert. Quantitative Methods: Motion Analysis, Audio Analysis, and Continuous Response Techniques. In Dorottya Fabian, Renee Timmers, and Emery Schubert, editors, *Expressiveness in music performance*, number July, pages 220–239. Oxford University Press, July 2014. ISBN 9780199659647. doi: 10.1093/acprof:oso/9780199659647.003.0013.

S Goldin-Meadow. Beyond words: the importance of gesture to researchers and learners. *Child development*, 71(1):231–9, 2000. ISSN 0009-3920. URL `http://www.ncbi.nlm.nih.gov/pubmed/10836578`.

Tobias Grosshauser and Thomas Hermann. Augmented haptics - An interactive feedback system for musicians. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5763 LNCS:100–108, 2009. ISSN 03029743. doi: 10.1007/978-3-642-04076-4{_}11.

Jonathan Grudin. A Moving Target: The evolution of human-computer interaction. In Julie A Jacko, editor, *HumanComputer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, pages 1–24. CRC Press, third edition, 2012. ISBN 9781439829431. doi: 10.1201/b11963-1. URL `http://research.microsoft.com/en-us/UM/People/jgrudin/publications/history/HCIhandbook3rd.pdf`.

J.P. Guilford. *Traits of Creativity.* McGraw-Hill, New York, NY, USA, 1967.

Hatice Gunes and Björn Schuller. Categorical and Dimensional Affect Analysis in Continuous Input: Current Trends and Future Directions. *Image and Vision Computing*, 31(2):120–136, July 2013. ISSN 02628856. doi: 10.1016/j.imavis.2012.06.016. URL `http://linkinghub.elsevier.com/retrieve/pii/S0262885612001084`.

Su Yeong Gwon, Chul Woo Cho, Hyeon Chang Lee, Won Oh Lee, and Kang Ryoung Park. Robust Eye and Pupil Detection Method for Gaze Tracking. *International Journal of Advanced Robotic Systems*, 10:1–7, 2013. ISSN 1729-8806. doi: 10.5772/55520. URL `http://www.intechopen.com/journals/international{_}journal{_}of{_}advanced{_}robotic{_}systems/robust-eye-and-pupil-detection-method-for-gaze-tracking`.

Aristotelis Hadjakos, Tobias Groß hauser, and Werner Goebl. Motion Analysis of Music Ensembles with the Kinect. In *International Conference on New Interfaces for Musical Expression (NIME)*, pages 106–110, KAIST, Daejeon, Korea, 2013.

L. Harmat, F. Ullen, O. de Manzano, E. Olsson, U. Elofsson, B. von Scheele, and T. Theorell. Heart Rate Variability During Piano Playing: A Case Study of Three Professional Solo Pianists Playing a Self-Selected and a Difficult Prima Vista Piece. *Music and Medicine*, 3(2):102–107, November 2010. ISSN 1943-8621. doi: 10.1177/1943862110387158.

Patrick G T Healey, Joe Leach, and Nick Bryan-kinns. Inter-Play : Understanding Group Music Improvisation as a Form of Everyday Interaction. In *Proceedings of Less is More: Simple Computing in an Age of Complexity*, 2005.

B A Hennessey, T M Amabile, and J S Mueller. Consensual Assessment. In M.A. Runco and S.R. Pritzker, editors, *Encyclopedia of Creativity*, volume 1, pages 253–260. Elsevier, 2 edition, 2011. ISBN 9780123750389. doi: 10.1016/B978-0-12-375038-9.00046-7.

Javier Hernandez, Daniel McDuff, Xavier Benavides, Judith Amores, Pattie Maes, and Rosalind Picard. AutoEmotive : Bringing Empathy to the Driving Experience to Manage Stress. *Proc. Designing Interactive Systems - DIS Companion*, pages 1–4, 2014. doi: 10.1145/2598784.2602780.

Nis Hjortskov, Dag Rissén, Anne Katrine Blangsted, Nils Fallentin, Ulf Lundberg, and Karen Sø gaard. The effect of mental stress on heart rate variability and blood pressure during computer work. *European journal of applied physiology*, 92(1-2): 84–9, June 2004. ISSN 1439-6319. doi: 10.1007/s00421-004-1055-z. URL `http://www.ncbi.nlm.nih.gov/pubmed/14991326`.

James Hollan, Edwin Hutchins, and David Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2):174–196, June 2000. ISSN 10730516. doi: 10.1145/353485.353487. URL `http://dl.acm.org/citation.cfm?id=353485.353487`.

Horst Holstein. Recognition of human periodic motion-a frequency domain approach. In *Object recognition supported by user interaction for service robots*, volume 1, pages 311–314. IEEE Comput. Soc, 2002. ISBN 0-7695-1695-X. doi: 10.1109/ICPR.2002.1044700. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1044700`.

Anthony J Hornof. The Prospects For Eye-Controlled Musical Performance. In *Proceedings of NIME 2014: The International Conference on New Interfaces for Musical Expression*, pages 461–466, 2014.

Dennis Howitt and Duncan Cramer. *Introduction to Research Methods in Psychology.* Pearson, 3 edition, 2011. ISBN 9780273734994.

Hayley Hung, Gwenn Englebienne, and Laura Cabrera Quiros. Detecting conversing groups with a single worn accelerometer. In *Proceedings of the 16^{th} International Conference on Multimodal Interaction - ICMI '14*, pages 84–91, New York, New York, USA, 2014. ACM Press. ISBN 9781450328852. doi: 10.1145/2663204.2663228. URL `http://dl.acm.org/citation.cfm?doid=2663204.2663228`.

Robert J K Jacob and Keith S Karn. Eye Tracking in Human Computer Interaction and Usability Research : Ready to Deliver the Promises. In Ralph Radach, Jukka Hyona, and Heiner Deubel, editors, *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, pages 573–605. Elsevier, 2003. ISBN 978-0444510204.

Robert G. Jaeger and Tim R. Halliday. On Confirmatory versus Exploratory Research. *Herpetologica*, 54(3):65–66, 1998.

Kay Redfield Jamison. *Touched with Fire: Manic-Depressive Illness and the Artistic Temperament.* Simon & Schuster Ltd, December 1996.

Joris H. Janssen, Wijnand a. IJsselsteijn, and Joyce H.D.M. Westerink. How affective technologies can influence intimate interactions and improve social connectedness. *International Journal of Human-Computer Studies*, 72(1):33–43, January 2014. ISSN 10715819. doi: 10.1016/j.ijhcs.2013.09.007. URL `http://linkinghub.elsevier.com/retrieve/pii/S1071581913001316`.

Natasha Jaques, Sara Taylor, Asaph Azaria, Asma Ghandeharioun, Akane Sano, and Rosalind Picard. Predicting students' happiness from physiology, phone, mobility, and behavioral data. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 222–228. IEEE, September 2015. ISBN 978-1-4799-9953-8. doi: 10.1109/ACII.2015.7344575. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7344575`.

Amir-Homayoun Javadi, Zahra Hakimi, Morteza Barati, Vincent Walsh, and Lili Tcheang. SET: a pupil detection method using sinusoidal approximation. *Frontiers in Neuroengineering*, 8(April):1–10, 2015. ISSN 1662-6443. doi: 10.3389/fneng.2015.00004.

Karl Kristoffer Jensen and Emanuela Marchetti. A Meta-study of musicians' non-verbal interaction. *The international journal of technology knowledge & society*, 6(5):1–12, 2010.

Alexander Refsum Jensenius, Marcelo M Wanderley, Rolf Inge Godøy, and Marc Leman. Musical Gestures: Concepts and Methods in Research. In Rolf Inge Godøy and Marc Leman, editors, *Musical Gestures: Sound, Movement, and Meaning*, chapter 2, pages 12–35. Routledge, 2010. ISBN 0415998867.

Patrick Jermann, Marc-Antoine Nüssli, and Weifeng Li. Using dual eye-tracking to unveil coordination and expertise in collaborative Tetris. In *BCS '10 Proceedings of the 24^{th} BCS Interaction Specialist Group Conference*, pages 36–44, Swinton, UK, 2010. British Computer Society.

Brendan John, Srinivas Sridharan, and Reynold Bailey. Collaborative eye tracking for image analysis. In *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '14*, pages 239–242, New York, New York, USA, 2014. ACM Press. ISBN 9781450327510. doi: 10.1145/2578153.2578215. URL `http://dl.acm.org/citation.cfm?doid=2578153.2578215`.

Kristiina Jokinen, Masafumi Nishida, and Seiichi Yamamoto. On eye-gaze and turn-taking. *Proceedings of the 2010 workshop on Eye gaze in intelligent human machine interaction - EGIHMI '10*, pages 118–123, 2010. doi: 10.1145/2002333.2002352. URL `http://portal.acm.org/citation.cfm?doid=2002333.2002352`.

Peter K Kaiser. Physiological response to color: A critical review. *Color Research & Application*, 9(1):29–36, March 1984. ISSN 1520-6378. doi: 10.1002/col.5080090106.

Michelle Karg, Kolja Kühnlenz, and Martin Buss. Recognition of affect based on gait patterns. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 40(4):1050–1061, 2010. ISSN 10834419. doi: 10.1109/TSMCB.2010.2044040.

Moritz Kassner, William Patera, and Andreas Bulling. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. *CoRR*, abs/1405.0, April 2014. URL `http://arxiv.org/abs/1405.0006`.

Naz Kaya and H. Epps. Color-emotion associations: Past experience and personal preference. *AIC 2004 Color and Paints, Interim Meeting of the International Color Association*, pages 31–34, 2004.

Johannes Keller, Herbert Bless, Frederik Blomann, and Dieter Kleinböhl. Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology*, 47(4):849–852, July 2011. ISSN 00221031. doi: 10.1016/j.jesp.2011.02.004. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022103111000321`.

Peter E Keller, D Fabian, R Timmers, and E Schubert. Ensemble performance: interpersonal alignment of musical expression. *Expressiveness in music performance: Empirical approaches across styles and cultures*, pages 260–282, 2014.

Adam Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22–63, January 1967. ISSN 00016918. doi: 10.1016/0001-6918(67)90005-4. URL `http://linkinghub.elsevier.com/retrieve/pii/0001691867900054`.

Adam Kendon. Spacing and Orientation in Co-present Interaction. *Lecture Notes in Computer Science*, 5967:1–15, 2010.

Barry J. Kenny and Martin Gellrich. Improvisation. In Richard Parncutt and Gary McPherson, editors, *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning*, chapter 8, pages 117–134. Oxford University Press, April 2002. ISBN 9780195138108. doi: 10.1093/acprof:oso/9780195138108.001.0001.

David A. Kenny, Deborah A. Kashy, and William L. Cook. *Dyadic Data Analysis*. Guilford Press, 2006. ISBN 9781572309869. URL `http://www.guilford.com/books/Dyadic-Data-Analysis/Kenny-Kashy-Cook/9781572309869`.

Raymond A Kent. *Analysing Quantitative Data: Variable-based and Case-based Approaches to Non-experimental Datasets*. SAGE Publications, 2015. ISBN 1473917913. URL `https://uk.sagepub.com/en-gb/eur/analysing-quantitative-data/book241117`.

Nanda Khaorapapong and Matthew Purver. Designing Unobtrusive Interfaces to Increase Naturalness of First Time Face-To-Face Interaction. In *5th International Workshop on Semantic Ambient Media Experience (SAME 2012) Special Theme 2012 Unobtrusive Semantic Ambient Media Beyond the Screen and Pixels"*, number Cmc, Newcastle, UK, 2012. URL `http://www.eecs.qmul.ac.uk/~mpurver/papers/khaorapapong-purver12same.pdf`.

Jonghwa Kim and Elisabeth André. Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(12):2067–83, December 2008. ISSN 0162-8828. doi: 10.1109/TPAMI.2008.26. URL http://www.ncbi.nlm.nih.gov/pubmed/18988943.

K. H. Kim, S. W. Bang, and S. R. Kim. Emotion recognition system using short-term monitoring of physiological signals. *Medical & Biological Engineering & Computing*, 42(3):419–427, May 2004. ISSN 0140-0118. doi: 10.1007/BF02344719.

Heinz K Klein and Michael D Myers. A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. *MIS Quarterly*, 23(1):67, March 1999. ISSN 02767783. doi: 10.2307/249410. URL http://www.jstor.org/stable/249410?origin=crossref.

Olivier Klein, Stéphane Doyen, Christophe Leys, Pedro a Magalhães de Saldanha da Gama, Sarah Miller, Laurence Questienne, and Axel Cleeremans. Low Hopes, High Expectations: Expectancy Effects and the Replicability of Behavioral Experiments. *Perspectives on Psychological Science*, 7(6):572–584, 2012. ISSN 1745-6916. doi: 10.1177/1745691612463704. URL http://pps.sagepub.com/content/7/6/572.full.pdf.

Chris L Kleinke. Gaze and eye contact: a research review. *Psychological bulletin*, 100(1):78–100, July 1986. ISSN 0033-2909. URL http://www.ncbi.nlm.nih.gov/pubmed/3526377.

Andrea Kleinsmith and Nadia Bianchi-Berthouze. Affective Body Expression Perception and Recognition: A Survey. *IEEE Transactions on Affective Computing*, 4(1):15–33, January 2013. ISSN 1949-3045. doi: 10.1109/T-AFFC.2012.16. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6212434.

R. Benjamin Knapp, Javier Jaimovich, and Niall Coghlan. Measurement of motion and emotion during musical performance. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–5. IEEE, September 2009. ISBN 978-1-4244-4800-5. doi: 10.1109/ACII.2009.5349469. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5349469.

Jacqueline Kory and Sidney D'Mello. Affect Elicitation for Affective Computing. In Rafael Calvo, Sidney D'Mello, Jonathan Gratch, and Arvid Kappas, editors, *The Oxford Handbook of Affective Computing*, chapter 27, pages 371–383. Oxford University Press, April 2014. doi: 10.1093/oxfordhb/9780199942237.013.001.

Irene Kotsia, Ioan Buciu, and Ioannis Pitas. An analysis of facial expression recognition under partial facial image occlusion. *Image and Vision Computing*, 26(7):1052–1067,

2008. ISSN 02628856. doi: 10.1016/j.imavis.2007.11.004. URL `http://linkinghub.elsevier.com/retrieve/pii/S0262885607002107`.

Yasuo Kuhara and Daiki Kobayashi. Kinetic Particles Synthesizer Using Multi-Touch Screen Interface of Mobile Devices. In *New Interfaces for Musical Expression (NIME11)*, number June, pages 136–137, 2011.

D. Kulic and E.A. Croft. Affective State Estimation for Human-Robot Interaction. *IEEE Transactions on Robotics*, 23(5):991–1000, October 2007. ISSN 1552-3098. doi: 10.1109/TRO.2007.904899. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4339537`.

Martin Kusserow, Gerhard Tröster, Victor Candia, Gerd Folkers, Oliver Amft, and Horst Hildebrandt. Wearable monitoring of stage fright in professional musicians. *Proceedings - International Symposium on Wearable Computers, ISWC*, 2010. ISSN 15504816. doi: 10.1109/ISWC.2010.5665852.

Beatrice C Lacey and John I Lacey. Studies of heart rate and other bodily processes in sensorimotor behavior. In *Cardiovascular psychophysiology: Current issues in response mechanisms, biofeedback and methodology*, pages 538–564. AldineTransaction, New Brunswick, NJ, US, 1974.

Stephen D Lapan, MaryLynn T Quartaroli, and Frances J Riemer. *Qualitative research: an introduction to methods and designs*. Jossey-Bass, 2011. ISBN 9780470548004.

Po-ming Lee, Chia-wei Chang, and Tzu-chien Hsiao. Can human decisions be predicted through heart rate changes? In *2010 Second World Congress on Nature and Biologically Inspired Computing (NaBIC)*, pages 189–193. IEEE, December 2010. ISBN 978-1-4244-7377-9. doi: 10.1109/NABIC.2010.5716336. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5716336`.

Wen-Yuan Lee, M. Ronnier Luo, and Li-Chen Ou. Assessing the affective feelings of two- and three-dimensional objects. *Color Research & Application*, 34(1):75–83, February 2009. ISSN 03612317. doi: 10.1002/col.20464.

María J Leone, Agustín Petroni, Diego Fernandez Slezak, and Mariano Sigman. The tell-tale heart: heart rate fluctuations index objective and subjective events during a game of chess. *Frontiers in human neuroscience*, 6 (October):273, January 2012. ISSN 1662-5161. doi: 10.3389/fnhum.2012.00273. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3465955&tool=pmcentrez&rendertype=abstract`.

Marion H. Levine and Brian Sutton-Smith. Effects of age, sex, and task on visual behavior during dyadic interaction. *Developmental Psychology*, 9(3):400–405, 1973. ISSN 1939-0599(Electronic);0012-1649(Print). doi: 10.1037/h0034929.

Chern Hong Lim, Ekta Vats, and Chee Seng Chan. Fuzzy human motion analysis: A review. *Pattern Recognition*, 48(5):1773–1796, 2014. ISSN 00313203. doi: 10.1016/j.patcog.2014.11.016. URL http://www.sciencedirect.com/science/article/pii/S0031320314004865.

Charles J Limb and Allen R Braun. Neural substrates of spontaneous musical performance: an FMRI study of jazz improvisation. *PloS one*, 3(2), January 2008. ISSN 1932-6203. doi: 10.1371/journal.pone.0001679. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2244806&tool=pmcentrez&rendertype=abstract.

Eric Limer. Apple Watch: Everything You Need to Know, 2014. URL http://gizmodo.com/apple-watch-everything-you-need-to-know-1632172509.

L.-O. Lundqvist, F. Carlsson, P. Hilmersson, and P. N. Juslin. Emotional responses to music: experience, expression, and physiology. *Psychology of Music*, 37(1):61–90, August 2008. ISSN 0305-7356. doi: 10.1177/0305735607086048.

Clemens Maidhof, Torsten Kästner, and Tommi Makkonen. Combining EEG, MIDI, and motion capture techniques for investigating musical performance. *Behavior research methods*, 46(1):185–95, 2014. ISSN 1554-3528. doi: 10.3758/s13428-013-0363-9. URL http://www.ncbi.nlm.nih.gov/pubmed/23943580.

Marek Malik, J. Thomas Bigger, A. John Camm, Robert E. Kleiger, Alberto Malliani, Arthur J. Moss, and Peter J. Schwartz. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17(3):354–81, March 1996. ISSN 0195-668X. URL http://eurheartj.oxfordjournals.org/content/17/3/354.extract.

Regan L. Mandryk and M. Stella Atkins. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, 65(4):329–347, April 2007. ISSN 10715819. doi: 10.1016/j.ijhcs.2006.11.011. URL http://linkinghub.elsevier.com/retrieve/pii/S1071581906001881.

Regan L. Mandryk, Kori M. Inkpen, and Thomas W. Calvert. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, 25(2):141–158, March 2006. ISSN 0144-929X. doi: 10.1080/01449290500331156.

Christopher D. Manning and Hinrich Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999. ISBN 0262133601. URL `https://mitpress.mit.edu/books/foundations-statistical-natural-language-processing`.

A Mazur, E Rosa, M Faupel, J Heller, R Leen, and B Thurman. Physiological aspects of communication via mutual gaze. *American Journal of Sociology*, 86(1):50–74, July 1980. ISSN 0002-9602. URL `http://www.ncbi.nlm.nih.gov/pubmed/7435770`.

Sebastián Mealla, A Väljamäe, M Bosi, and S Jordà. Listening to Your Brain: Implicit Interaction in Collaborative Music Performances. In *The International Conference on New Interfaces for Musical Expression (NIME)*, number June, pages 149–154, Oslo, Norway, 2011. URL `http://mtg.upf.es/system/files/publications/listening_to_your_brain_camera_ready.pdf`.

Carolina Brum Medeiros and Marcelo M. Wanderley. A comprehensive review of sensors and instrumentation methods in devices for musical expression. *Sensors (Switzerland)*, 14(8):13556–13591, 2014. ISSN 14248220. doi: 10.3390/s140813556.

S a Mednick. The associative basis of the creative process. *Psychological review*, 69(3): 220–32, May 1962. ISSN 0033-295X. URL `http://www.ncbi.nlm.nih.gov/pubmed/14472013`.

Angeliki Metallinou, Athanassios Katsamanis, Yun Wang, and Shrikanth Narayanan. Tracking changes in continuous emotion states using body language and prosodic cues. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, volume 2, pages 2288–2291. IEEE, May 2011. ISBN 978-1-4577-0538-0. doi: 10.1109/ICASSP.2011.5946939. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5946939`.

Angeliki Metallinou, Athanasios Katsamanis, and Shrikanth Narayanan. Tracking continuous emotional trends of participants during affective dyadic interactions using body language and speech information. *Image and Vision Computing*, 31 (2):137–152, 2013. ISSN 0262-8856. doi: 10.1016/j.imavis.2012.08.018. URL `http://www.sciencedirect.com/science/article/pii/S0262885612001710`.

J E Mietus, C-K Peng, I Henry, R L Goldsmith, and a L Goldberger. The pNNx files: re-examining a widely used heart rate variability measure. *Heart (British Cardiac Society)*, 88(4):378–80, October 2002. ISSN 1468-201X. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1767394&tool=pmcentrez&rendertype=abstract.

Dan Misener. Apple buys Emotient emotion detection company, 2016. URL http://www.cbc.ca/news/technology/apple-emotiont-1.3400458.

Thomas Mitchell, Sebastian Madgwick, and Imogen Heap. Musical Interaction with Hand Posture and Orientation : A Toolbox of Gestural Control Mechanisms. In *New Interfaces for Musical Expression (NIME12)*, 2012.

Ingrid Monson. *Saying Something: Jazz Improvisation and Interaction*. Chicago Studies in Ethnomusicology. University of Chicago Press, Chicago, IL, 1996. ISBN 9780226534770. URL http://press.uchicago.edu/ucp/books/book/chicago/S/bo3621755.html.

Evan Morgan and Hatice Gunes. Human nonverbal behaviour understanding in the wild for new media art. In AlbertAli Salah, Hayley Hung, Oya Aran, and Hatice Gunes, editors, *Human Behavior Understanding*, volume 8212 of *Lecture Notes in Computer Science*, pages 27–39. Springer International Publishing, 2013. ISBN 978-3-319-02713-5. doi: 10.1007/978-3-319-02714-2_3.

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. Measuring affect for the study and enhancement of co-present creative collaboration. In *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*, pages 659–664, 2013. ISBN 9780769550480. doi: 10.1109/ACII.2013.115. URL http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6681506.

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. Instrumenting the Interaction : Affective and psychophysiological features of live collaborative musical improvisation. In *14th International Conference on New Interfaces for Musical Expression (NIME14)*, London, 2014. URL http://www.nime.org/proceedings/2014/nime2014_353.pdf.

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. Using affective and behavioural sensors to explore aspects of collaborative music making. *International Journal of Human-Computer Studies*, 82:31–47, 2015a. ISSN 1071-5819. doi: 10.1016/j.

ijhcs.2015.05.002. URL http://www.sciencedirect.com/science/article/pii/S1071581915000853.

Evan Morgan, Hatice Gunes, and Nick Bryan-Kinns. The LuminUs: Providing musicians with visual feedback on the gaze and body motion of their co-performers. In Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction INTERACT 2015*, volume 9297 of *Lecture Notes in Computer Science*, pages 47–54. Springer International Publishing, 2015b. ISBN 978-3-319-22667-5. doi: 10.1007/978-3-319-22668-2_4.

Romy Müller, Jens R Helmert, Sebastian Pannasch, and Boris M Velichkovsky. Gaze transfer in remote cooperation: is it always helpful to see what your partner is attending to? *Quarterly journal of experimental psychology*, 66(7):1302–16, January 2012. ISSN 1747-0226. doi: 10.1080/17470218.2012.737813. URL http://www.ncbi.nlm.nih.gov/pubmed/23140500.

Romy Müller, Jens R Helmert, Sebastian Pannasch, and Boris M Velichkovsky. Gaze transfer in remote cooperation: is it always helpful to see what your partner is attending to? *Quarterly journal of experimental psychology*, 66(7):1302–16, January 2012. ISSN 1747-0226. doi: 10.1080/17470218.2012.737813. URL http://www.ncbi.nlm.nih.gov/pubmed/23140500.

Fiona Mulvey. Woking copy of definitions and terminology for eye tracker accuracy and precision, 2010.

Hidehiro Nakahara, Shinichi Furuya, Satoshi Obata, Tsutomu Masuko, and Hiroshi Kinoshita. Emotion-related changes in heart rate and its variability during performance and perception of music. *Annals of the New York Academy of Sciences*, 1169: 359–62, July 2009. ISSN 1749-6632. doi: 10.1111/j.1749-6632.2009.04788.x. URL http://www.ncbi.nlm.nih.gov/pubmed/19673808.

Fatma Nasoz, Kaye Alvarez, Christine L. Lisetti, and Neal Finkelstein. Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cognition, Technology & Work*, 6(1):4–14, February 2004. ISSN 1435-5558. doi: 10.1007/s10111-003-0143-x.

James Nazroo and Sue Arthur. Designing Fieldwork Strategies and Materials. In Jane Ritchie and Jane Lewis, editors, *Qualitative research practice: A guide for social science students and researchers*, chapter 5, pages 109–137. SAGE Publications, London, 2003.

Hartmut Neven. Gaze tracking system, 2013.

K.A. Nichols and B.G. Champness. Eye gaze and the GSR. *Journal of Experimental Social Psychology*, 7(6):623–626, November 1971. ISSN 00221031. doi: 10.1016/0022-1031(71)90024-2. URL `http://linkinghub.elsevier.com/retrieve/pii/0022103171900242`.

Lena Nowicki, Wolfgang Prinz, Marc Grosjean, Bruno H. Repp, and Peter E. Keller. Mutual adaptive timing in interpersonal action coordination. *Psychomusicology: Music, Mind, and Brain*, 23(1):6–20, 2013. ISSN 2162-1535. doi: 10.1037/a0032039.

Marc-Antoine Nüssli. *Dual Eye-Tracking Methods for the Study of Remote Collaborative Problem Solving*. PhD thesis, Swiss Federal Institute of Technology, Lausanne, 2011.

Marc-Antoine Nüssli and Patrick Jermann. Effects of sharing text selections on gaze cross-recurrence and interaction quality in a pair programming task. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, page 1125, 2012. doi: 10.1145/2145204.2145371. URL `http://dl.acm.org/citation.cfm?doid=2145204.2145371`.

Z. Obrenovic and D. Starcevic. Modeling multimodal human-computer interaction. *Computer*, 37(9):65–72, September 2004. ISSN 0018-9162. doi: 10.1109/MC.2004.139. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1333009`.

M Oehler, V Ling, K Melhorn, and M Schilling. A multichannel portable ECG system with capacitive sensors. *Physiological measurement*, 29(7):783–93, July 2008. ISSN 0967-3334. doi: 10.1088/0967-3334/29/7/007. URL `http://www.ncbi.nlm.nih.gov/pubmed/18560053`.

Catharine Oertel, Marcin Wlodarczak, Jens Edlund, Petra Wagner, and Joakim Gustafson. Gaze Patterns in Turn-Taking. In *Proceedings of Interspeech*, volume 29, February 2012.

Martin T. Orne. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11):776–783, 1962. ISSN 0003-066X. doi: 10.1037/h0043424.

Li-Chen Ou, M. Ronnier Luo, Andre Woodcock, and Angela Wright. A study of colour emotion and colour preference. Part I: Colour emotions for single colours.

Color Research & Application, 29(3):232–240, June 2004. ISSN 0361-2317. doi: 10.1002/col.20010.

Maja Pantic, Roderick Cowie, Francesca D'Errico, Dirk Heylen, Marc Mehu, Catherine Pelachaud, Isabella Poggi, Marc Schroeder, and Alessandro Vinciarelli. Social signal processing: The research agenda. *Visual Analysis of Humans*, pages 511–538, 2011. doi: 10.1007/978-0-85729-997.

M. Patel, S.K.L. Lal, D. Kavanagh, and P. Rossiter. Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications*, 38(6):7235–7242, June 2011. ISSN 09574174. doi: 10.1016/j.eswa.2010.12.028. URL `http://linkinghub.elsevier.com/retrieve/pii/S0957417410013916`.

Miles L. Patterson. A sequential functional model of nonverbal exchange. *Psychological Review*, 89(3):231–249, 1982. ISSN 0033-295X. doi: 10.1037/0033-295X.89.3.231. URL `http://content.apa.org/journals/rev/89/3/231`.

Marcus Thomas Pearce. *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition.* PhD thesis, City University, London, 2005.

Corinna Peifer. Advances in Flow Research. pages 139–164, 2012. doi: 10.1007/978-1-4614-2359-1.

A. Pentland. Socially aware, computation and communication. *Computer*, 38(3):33–40, March 2005. ISSN 0018-9162. doi: 10.1109/MC.2005.104. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1413116`.

Ulrich J Pfeiffer, Kai Vogeley, and Leonhard Schilbach. From gaze cueing to dual eye-tracking: novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience and biobehavioral reviews*, 37(10 Pt 2):2516–28, December 2013. ISSN 1873-7528. doi: 10.1016/j.neubiorev.2013.07.017. URL `http://www.ncbi.nlm.nih.gov/pubmed/23928088`.

Dung Phan, Lee Yee Siong, Pubudu N Pathirana, and Aruna Seneviratne. Smartwatch: Performance evaluation for long-term heart rate monitoring. In *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, pages 144–147. IEEE, October 2015. ISBN 978-1-4673-6609-0. doi: 10.1109/ISBB.2015.7344944. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7344944`.

Philips Electronics. Philips Vital Signs Camera - Measurement Tips, 2014. URL http://www.vitalsignscamera.com/measurement-tips.html.

Rosalind W. Picard. Affective computing. Technical report, Cambridge, MA, USA, October 1997. URL http://dl.acm.org/citation.cfm?id=265013.

Rosalind W. Picard. *Affective Computing*. The MIT Press, 2000. ISBN 0262661152.

RW Picard, Elias Vyzas, and Jennifer Healey. Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1175–1191, 2001. ISSN 01628828. doi: 10.1109/34.954607. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=954607http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=954607.

M Pietrowicz and K Karahalios. Sonic Shapes: Visualizing Vocal Expression. In *The 19th International Conference on Auditory Display (ICAD-2013)*, 2013. URL http://social.cs.uiuc.edu/papers/pdfs/13-pietrowicz-icad.pdf.

Plessey Semiconductors. Non-contact ECG measurement using EPIC. Technical Report 1, Swindon, 2012.

Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE transactions on bio-medical engineering*, 58(1):7–11, January 2011. ISSN 1558-2531. doi: 10.1109/TBME.2010.2086456. URL http://www.ncbi.nlm.nih.gov/pubmed/20952328.

F E Pollick, H M Paterson, A Bruderlin, and a J Sanford. Perceiving affect from arm movement. *Cognition*, 82(2):B51–61, December 2001. ISSN 0010-0277. URL http://www.ncbi.nlm.nih.gov/pubmed/11716834.

Keith S. Potter, Geraint A. Wiggins, and Marcus T. Pearce. Towards Greater Objectivity in Music Theory: Information-Dynamic Analysis of Minimalist Music. XI:295–324, 2007. ISSN 10298649. doi: 10.1177/102986490701100207. URL http://eprints.gold.ac.uk/998/.

Jorjanna Price. Plug in to movies that know how you're feeling. Online, March 2011. URL http://www.newscientist.com/blogs/culturelab/2011/03/plug-in-for-emotional-response-cinema.html. (Accessed 10th December 2012).

Emily Pronin. The Introspection Illusion. In *Advances in Experimental Social Psychology*, volume 41, chapter 1, pages 1–67. Elsevier Inc., 2009. ISBN 9780123744722. doi: 10.1016/S0065-2601(08)00401-2.

T Pursche, J Krajewski, and Reinhard Moeller. Video-based heart rate measurement from human faces. In *2012 IEEE International Conference on Consumer Electronics (ICCE)*, pages 544–545. IEEE, January 2012. ISBN 978-1-4577-0231-0. doi: 10.1109/ICCE.2012.6161965. URL `http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=6161965`.

Luis Quesada and Alejandro León. 3D Markerless Motion Tracking in Real-Time Using a Single Camera. In *Intelligent Data Engineering and Automated Learning - IDEAL 2011*, pages 186–193. Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-23878-9{_}23.

Rafael Ramirez and Zacharias Vamvakousis. Detecting emotion from EEG signals using the Emotive Epoc device. In Fabio Massimo Zanzotto, Shusaku Tsumoto, Niels Taatgen, and Yiyu Yao, editors, *Brain Informatics*, volume 7670 LNAI, pages 175–184. Springer Berlin Heidelberg, 2012. ISBN 9783642351389. doi: 10.1007/978-3-642-35139-6{_}17.

John T. Ramshur. *Design, Evaluation, and Application of Heart Rate Variability Software (HRVAS)*. PhD thesis, The University of Memphis, 2010.

Dave Randall and Mark Rouncefield. Ethnography. In Mads Soegaard and Rikke Friis Dam, editors, *The Encyclopedia of Human-Computer Interaction*, chapter 31. The Interaction Design Foundation, Aarhus, Denmark, 2 edition, 2013.

Niklas Ravaja, Timo Saari, Mikko Salminen, Jari Laarni, and Kari Kallinen. Phasic Emotional Reactions to Video Game Events: A Psychophysiological Investigation. *Media Psychology*, 8(4):343–367, November 2006. ISSN 1521-3269. doi: 10.1207/s1532785xmep0804_2.

Rebecca G Reed, Ashley K Randall, Jessica H Post, and Emily a Butler. Partner influence and in-phase versus anti-phase physiological linkage in romantic couples. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, 88(3):309–16, June 2013. ISSN 1872-7697. doi: 10.1016/j.ijpsycho.2012.08.009. URL `http://www.ncbi.nlm.nih.gov/pubmed/22922526`.

Leah M. Reeves, Jean-Claude Martin, Michael McTear, Tv Raman, Kay M. Stanney, Hui Su, Qian Ying Wang, Jennifer Lai, James a. Larson, Sharon Oviatt, T. S. Balaji, Stéphanie Buisine, Penny Collings, Phil Cohen, and Ben Kraal. Guidelines for multimodal user interface design. *Communications of the ACM*, 47(1):57, 2004. ISSN 00010782. doi: 10.1145/962081.962106.

Dennis Reidsma, Mustafa Radha, and Anton Nijholt. Mediated Interactions and Musical Expression - A Survey. In Newton Lee, editor, *Digital Da Vinci: Computers in Music*, pages 79–98. Springer, New York, NY, USA, 2014. ISBN 9781493905355. doi: 10.1007/978-1-4939-0536-2_4.

Bernd Reiter. The Epistemology and Methodology of Exploratory Social Science Research : Crossing Popper with Marcuse. *The Dialectics of Citizenship: Exploring Privilege, Exclusion, and Racialization*, pages 1–23, 2013.

Zhou Ren, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang. Robust part-based hand gesture recognition using kinect sensor. *IEEE Transactions on Multimedia*, 15 (5):1110–1120, 2013. ISSN 15209210. doi: 10.1109/TMM.2013.2246148.

Charles Rich, Brett Ponsler, Aaron Holroyd, and Candace L. Sidner. Recognizing engagement in human-robot interaction. *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 375–382, March 2010. doi: 10.1109/HRI.2010.5453163. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5453163.

Daniel C Richardson and Rick Dale. Looking to understand: the coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive science*, 29(6):1045–60, November 2005. ISSN 0364-0213. doi: 10.1207/s15516709cog0000_29. URL http://www.ncbi.nlm.nih.gov/pubmed/21702802.

Jane Ritchie. The Applications of Qualitative Methods to Social Research. In Jane Ritchie and Jane Lewis, editors, *Qualitative research practice: A guide for social science students and researchers*, chapter 2, pages 24–46. SAGE Publications, London, 2003.

Dw Rowe, John Sibert, and Don Irwin. Heart rate variability: Indicator of user state as an aid to human-computer interaction. *Proceedings of the SIGCHI conference on ...*, pages 480–487, 1998. doi: 10.1145/274644.274709. URL http://dl.acm.org/citation.cfm?id=274709.

Celementina R. Russo. The Quantified Self. In *Foundations of Augmented Cognition*, pages 514–520. 2015. doi: 10.1007/978-3-319-20816-9{_}49.

H Sacks, EA Schegloff, and G Jefferson. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4):696–735, 1974. doi: 10.2307/412243.

Johanna Sänger, Viktor Müller, and Ulman Lindenberger. Intra- and interbrain synchronization and network properties when playing guitar in duets. *Frontiers in human neuroscience*, 6(November):312, 2012. ISSN 1662-5161. doi: 10.3389/fnhum.2012.00312. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3509332{&}tool=pmcentrez{&}rendertype=abstract.

Jyotirmay Sanghvi, Ginevra Castellano, Iolanda Leite, André Pereira, Peter W. McOwan, and Ana Paiva. Automatic analysis of affective postures and body motion to detect engagement with a game companion. In *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, page 305, New York, New York, USA, 2011. ACM Press. ISBN 9781450305617. doi: 10.1145/1957656.1957781. URL http://portal.acm.org/citation.cfm?doid=1957656.1957781.

Evangelos Sariyanidi, Hatice Gunes, and Andrea Cavallaro. Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(6):1113–1133, June 2015. ISSN 0162-8828. doi: 10.1109/TPAMI.2014.2366127. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6940284.

R. Keith Sawyer. Music and Conversation. In Dorothy Miell, Raymond MacDonald, and David J. Hargreaves, editors, *Musical Communication*, chapter 3. Oxford University Press, July 2005. ISBN 9780198529361. doi: 10.1093/acprof:oso/9780198529361.001.0001.

Robert Keith Sawyer. *Group Creativity: Music, Theater, Collaboration*. Psychology Press, 2003. ISBN 0805844368.

Laela Sayigh, Nicola Quick, Gordon Hastie, and Peter Tyack. Repeated call types in short-finned pilot whales, Globicephala macrorhynchus. *Marine Mammal Science*, 29(2):312–324, 2013. ISSN 08240469. doi: 10.1111/j.1748-7692.2012.00577.x.

Andrea Schiavio and S. Hoffding. Playing together without communicating? A pre-reflective and enactive account of joint musical performance. *Musicae Scientiae*, 19 (4):366–388, December 2015. ISSN 1029-8649. doi: 10.1177/1029864915593333.

Louis A. Schmidt and Laurel J. Trainor. Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognition & Emotion*, 15(4): 487–500, July 2001. ISSN 0269-9931. doi: 10.1080/02699930126048.

Frank Schneider, Wolfgang Grodd, Ute Weiss, Uwe Klose, Katharina R Mayer, Thomas Nägele, and Ruben C Gur. Functional MRI reveals left amygdala activation during emotion. *Psychiatry Research: Neuroimaging*, 76(2-3):75–82, December 1997. ISSN 09254927. doi: 10.1016/S0925-4927(97)00063-2.

Michael F Schober and Neta Spiro. Jazz improvisers' shared understanding: a case study. *Frontiers in psychology*, 5(August):808, January 2014. ISSN 1664-1078. doi: 10.3389/fpsyg.2014.00808. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4126153&tool=pmcentrez&rendertype=abstract`.

Björn Schuller and Felix Weninger. Human Affect Recognition: Audio-Based Methods. In *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–13. John Wiley & Sons, Inc., Hoboken, NJ, USA, September 2015. doi: 10.1002/047134608X.W8271.

Alfred Schutz. Making music together. *Collected papers II*, 1976. doi: 10.1007/978-94-010-1340-6_8.

Andrew Sears and Julie A Jacko, editors. *Human Computer Interaction: Fundamentals*. CRC Press, 2007. ISBN 9781420088816.

Frederick a. Seddon. Modes of communication during jazz improvisation. *British Journal of Music Education*, 22(1):47–61, March 2005. ISSN 0265-0517. doi: 10.1017/S0265051704005984. URL `http://www.journals.cambridge.org/abstract_S0265051704005984`.

SensoMotoric Instruments (SMI). SMI Eye Tracking Glasses 2 Wireless, 2015. URL `http://www.eyetracking-glasses.com/products/eye-tracking-glasses-2-wireless/technology/`.

D K Simonton. Creativity, task complexity, and intuitive versus analytical problem solving. *Psychological reports*, 37(2):351–4, October 1975. ISSN 0033-2941. doi: 10.2466/pr0.1975.37.2.351. URL `http://www.ncbi.nlm.nih.gov/pubmed/1187912`.

Robert Simpson, Kevin R Page, and David De Roure. Zooniverse: observing the world's largest citizen science platform. *Proceedings of the 23rd International Conference on World Wide Web*, pages 1049–1054, 2014. doi: 10.1145/2567948.2579215. URL `http://dl.acm.org/citation.cfm?id=2579215papers3://publication/doi/10.1145/2567948.2579215`.

Gary W Small, Teena D Moody, Prabha Siddarth, and Susan Y Bookheimer. Your brain on Google: patterns of cerebral activation during internet searching. *The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry*, 17(2):116–26, February 2009. ISSN 1545-7214. doi: 10.1097/JGP.0b013e3181953a02. URL `http://www.ncbi.nlm.nih.gov/pubmed/19155745`.

Dawn Snape and Liz Spencer. The foundations of qualitative research. In Jane Ritchie and Jane Lewis, editors, *Qualitative research practice: A guide for social science students and researchers*, chapter 1, pages 1–23. SAGE Publications, London, 2003. ISBN 0 7619 7109 2.

R J M Somsen, M W van der Molen, and J F Orlebeke. Phasic Heart Rate Changes in Reaction Time, Shock Avoidance, and Unavoidable Shock Tasks: Are Hypothetical Generalizations About Different S1S2 Tasks Justified? *Psychophysiology*, 20(1): 88–94, January 1983. ISSN 1469-8986. doi: 10.1111/j.1469-8986.1983.tb00908.x.

Soundbrenner. Soundbrenner Pulse, 2016. URL `http://www.soundbrenner.com/`.

Robert A. Stebbins. *Exploratory Research in the Social Sciences*. SAGE Publications, 2001. ISBN 0761923993. URL `https://books.google.com/books?hl=en{&}lr={&}id=hDE13{_}a{_}oEsC{&}pgis=1`.

Lindsey Sterling, Geraldine Dawson, Sara Webb, Michael Murias, Jeffrey Munson, Heracles Panagiotides, and Elizabeth Aylward. The role of face familiarity in eye tracking of faces by individuals with autism spectrum disorders. *Journal of autism and developmental disorders*, 38(9):1666–75, October 2008. ISSN 0162-3257. doi: 10.1007/s10803-008-0550-1. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2989724&tool=pmcentrez&rendertype=abstract`.

D. Stowell, a. Robertson, N. Bryan-Kinns, and M.D. Plumbley. Evaluation of live humancomputer music-making: Quantitative and qualitative approaches. *International Journal of Human-Computer Studies*, 67(11):960–975, November 2009. ISSN 10715819. doi: 10.1016/j.ijhcs.2009.05.007. URL `http://linkinghub.elsevier.com/retrieve/pii/S107158190900069X`.

Bettina Studer and Luke Clark. Place your bets: psychophysiological correlates of decision-making under risk. *Cognitive, affective & behavioral neuroscience*, 11(2):144–58, June 2011. ISSN 1531-135X. doi: 10.3758/s13415-011-0025-2. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3084947&tool=pmcentrez&rendertype=abstract`.

Hyeon-Jeong Suk and Hans Irtel. Emotional response to color across media. *Color Research & Application*, 35(1):64–77, February 2010. ISSN 03612317. doi: 10.1002/col.20554.

John R Suler. Primary process thinking and creativity. *Psychological Bulletin*, 88 (1):144–165, 1980. ISSN 0033-2909. doi: 10.1037/0033-2909.88.1.144. URL `http://content.apa.org/journals/bul/88/1/144`.

Melanie Swan. The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. *Big Data*, 1(2):85–99, 2013. ISSN 2167-6461. doi: 10.1089/big.2012.0002.

G. S. Taylor and C. Schmidt. Empirical Evaluation of the Emotiv EPOC BCI Headset for the Detection of Mental Actions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1):193–197, 2012. ISSN 1071-1813. doi: 10.1177/1071181312561017.

Tobii Technology. Tobii Eye Tracking: An introduction to eye tracking and Tobii Eye Trackers. Technical report, Tobii Technology, 2010. URL `http://www.tobii.com/Global/Analysis/Training/WhitePapers/Tobii_EyeTracking_Introduction_WhitePaper.pdf?epslanguage=en`.

Sherry Turkle. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books, 2012. ISBN 0465031463.

SelinDeniz Uzunosmanolu and MuratPerit Çakir. Examining an Online Collaboration Learning Environment with the Dual Eye-Tracking Paradigm: The Case of Virtual Math Teams. In Panayiotis Zaphiris and Andri Ioannou, editors, *Learning and Collaboration Technologies. Designing and Developing Novel Learning Experiences SE - 44*, volume 8523 of *Lecture Notes in Computer Science*, pages 462–472. Springer International Publishing, 2014. ISBN 978-3-319-07481-8. doi: 10.1007/978-3-319-07482-5{_}44.

Janet van der Linden, Rose Johnson, Jonathan Bird, Yvonne Rogers, and Erwin Schoonderwaldt. Buzzing to play. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, page 533, New York, New York, USA, 2011. ACM Press. ISBN 9781450302289. doi: 10.1145/1978942.1979017. URL `http://dl.acm.org/citation.cfm?doid=1978942.1979017`.

M W Van der Molen, R J M Somsen, J R Jennings, and J F Orlebeke. Sensorimotor integration and the timing of heart rate deceleration. *Psychophysiology of cardiovascular control: Models, methods and data*, pages 565–582, 1985.

Giovanna Varni, Antonio Camurri, Paolo Coletta, and Gualtiero Volpe. Emotional entrainment in music performance. In *2008 8^{th} IEEE International Conference on Automatic Face & Gesture Recognition*, pages 1–5. IEEE, September 2008. ISBN 978-1-4244-2153-4. doi: 10.1109/AFGR.2008.4813433. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4813433http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4813433.

Bogdan Vera, Elaine Chew, and PGT Healey. A Study of Ensemble Performance Under Restricted Line of Sight. In *Proceedings of the International Conference on Music Information Retrieval*, Curitiba, Brazil, 2013. URL http://www.ppgia.pucpr.br/ismir2013/wp-content/uploads/2013/09/171_Paper.pdf.

Roel Vertegaal. The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99*, pages 294–301, New York, New York, USA, 1999. ACM Press. ISBN 0201485591. doi: 10.1145/302979.303065. URL http://portal.acm.org/citation.cfm?doid=302979.303065.

Roel Vertegaal, Robert Slagter, Gerrit van der Veer, and Anton Nijholt. Eye gaze patterns in conversations. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '01*, pages 301–308, New York, New York, USA, 2001. ACM Press. ISBN 1581133278. doi: 10.1145/365024.365119. URL http://portal.acm.org/citation.cfm?doid=365024.365119.

Alessandro Vinciarelli, Maja Pantic, and Hervé Bourlard. Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 27(12):1743–1759, November 2009. ISSN 02628856. doi: 10.1016/j.imavis.2008.11.007. URL http://linkinghub.elsevier.com/retrieve/pii/S0262885608002485.

Alessandro Vinciarelli, Maja Pantic, Dirk Heylen, Catherine Pelachaud, Isabella Poggi, F. D'Errico, and M. Schroeder. Bridging the Gap between Social Animal and Unsocial Machine: A Survey of Social Signal Processing. *IEEE Transactions on Affective Computing*, 3(1):69–87, January 2012. ISSN 1949-3045. doi: 10.1109/T-AFFC.2011.27. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5989788.

Gualtiero Volpe. *Computational Models of Expressive Gesture in Multimedia Systems.* PhD thesis, University of Genova, 2003.

Peter Votava and Erich Berger. The Heart Chamber Orchestra: An Audio-Visual Real-Time Performance For Chamber Orchestra Based On Heartbeats. *eContact!*, 14(2), 2012. URL `http://cec.sonus.ca/econtact/14_2/votava-berger_hco.html`.

Richard D Walk and Carolyn P Homan. Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society*, 22(5):437–440, November 1984. ISSN 0090-5054. doi: 10.3758/BF03333870.

Harald G Wallbott. Bodily expression of emotion. *European Journal of Social Psychology*, 28(6):879–896, November 1998. ISSN 0046-2772. doi: 10.1002/(SICI)1099-0992(1998110)28:6⟨879::AID-EJSP901⟩3.0.CO;2-W.

Marian Walter, Benjamin Eilebrecht, Tobias Wartzek, and Steffen Leonhardt. The smart car seat: personalized monitoring of vital signs in automotive applications. *Personal and Ubiquitous Computing*, 15(7):707–715, January 2011. ISSN 1617-4909. doi: 10.1007/s00779-010-0350-4.

Mark Weiser. The Computer for the 31st Century. *Scientific American*, 265(3):94–104, September 1991.

Mark Weiser. Hot topics-ubiquitous computing. *Computer*, 26(10):71–72, 1993. ISSN 00189162. doi: 10.1109/2.237456. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=237456`.

Mark Weiser. How computers will be used differently in the next twenty years. In *Proceedings of the 1999 IEEE Symposium on Security and Privacy (Cat. No.99CB36344)*, pages 234–235. Xerox PARC, IEEE Comput. Soc, 1999. ISBN 0-7695-0176-1. doi: 10.1109/SECPRI.1999.766917. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=766917`.

Brady T West, Kathleen B Welch, and Andrzej T Galecki. *Linear Mixed Models: A Practical Guide Using Statistical Software.* Chapman and Hall/CRC, 2007. ISBN 9781584884804.

Bruno Wicker, David I Perrett, Simon Baron-Cohen, and Jean Decety. Being the target of another's emotion: a PET study. *Neuropsychologia*, 41(2):139–46, January 2003. ISSN 0028-3932. URL `http://www.ncbi.nlm.nih.gov/pubmed/12459212`.

Matthias J Wieser, Paul Pauli, Georg W Alpers, and Andreas Mühlberger. Is eye to eye contact really threatening and avoided in social anxiety?–An eye-tracking and psychophysiology study. *Journal of anxiety disorders*, 23(1):93–103, January 2009. ISSN 0887-6185. doi: 10.1016/j.janxdis.2008.04.004. URL `http://www.ncbi.nlm.nih.gov/pubmed/18534814`.

Geraint A Wiggins, Marcus T Pearce, and M Daniel. Computational Modelling of Music Cognition and Musical Creativity. In Roger T. Dean, editor, *The Oxford Handbook of Computer Music*, pages 383–420. Oxford University Press, Oxford, 2009. ISBN 978-0-19-533161-5.

Aaron Williamon, Lisa Aufegger, and Hubert Eiholzer. Simulating and stimulating performance: introducing distributed simulation to enhance musical learning and performance. *Frontiers in Psychology*, 5(February):1–9, 2014. ISSN 1664-1078. doi: 10.3389/fpsyg.2014.00025.

Glen P. Williams and Chris L. Kleinke. Effects of Mutual Gaze and Touch on Attraction, Mood, and Cardiovascular Reactivity. *Journal of Research in Personality*, 27(2): 170–183, June 1993. ISSN 00926566. doi: 10.1006/jrpe.1993.1012. URL `http://linkinghub.elsevier.com/retrieve/pii/S0092656683710123`.

Andrew Wilson. Sensor- and Recognition-Based Input for Interaction. In Julie A Jacko, editor, *HumanComputer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, chapter 7, pages 133–156. CRC Press, third edition, 2012. doi: 10.1201/b11963-10.

G. B. Wilson and R. a. R. MacDonald. The sign of silence: Negotiating musical identities in an improvising ensemble. *Psychology of Music*, 40(5):558–573, August 2012. ISSN 0305-7356. doi: 10.1177/0305735612449506.

Andrea Stevenson Won, Jeremy N Bailenson, and Joris H Janssen. Automatic Detection of Nonverbal Behavior Predicts Learning in Dyadic Interactions. *IEEE Transactions on Affective Computing*, 5(2):112–125, April 2014. ISSN 1949-3045. doi: 10.1109/TAFFC.2014.2329304. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6827904`.

Robert Wood, Brian Maraj, C Matthew Lee, and Rafael Reyes. Short-term heart rate variability during a cognitive challenge in young and older adults. *Age and ageing*, 31(2):131–5, March 2002. ISSN 0002-0729. URL `http://www.ncbi.nlm.nih.gov/pubmed/11937476`.

Michaela B. Wooldridge and Jennifer Shapka. Playing with technology: Mothertoddler interaction scores lower during play with electronic toys. *Journal of Applied Developmental Psychology*, 33(5):211–218, September 2012. ISSN 01933973. doi: 10.1016/j.appdev.2012.05.005. URL `http://linkinghub.elsevier.com/retrieve/pii/S0193397312000445`.

A L Yarbus. *Eye movements and vision.* Plenum, New York, 1967.

Zhefan Ye, Yin Li, Alireza Fathi, Yi Han, Agata Rozga, Gregory D Abowd, and James M Rehg. Detecting eye contact using wearable eye-tracking glasses. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, page 699, New York, New York, USA, 2012. ACM Press. ISBN 9781450312240. doi: 10.1145/2370216.2370368. URL `http://dl.acm.org/citation.cfm?doid=2370216.2370368`.

Zhihong Zeng, Maja Pantic, Glenn I Roisman, and Thomas S Huang. A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):39–58, January 2009. ISSN 0162-8828. doi: 10.1109/TPAMI.2008.52. URL `http://www.ncbi.nlm.nih.gov/pubmed/19029545`.

Huiyu Zhou, Thomas Stone, Huosheng Hu, and Nigel Harris. Use of multiple wearable inertial sensors in upper limb motion tracking. *Medical engineering & physics*, 30 (1):123–33, January 2008. ISSN 1350-4533. doi: 10.1016/j.medengphy.2006.11.010. URL `http://www.ncbi.nlm.nih.gov/pubmed/17251049`.

Appendices

Appendix A

Study 1 Materials

A.1 Ethical Approval

Queen Mary
University of London

Queen Mary, University of London
Room W117
Queen's Building
Queen Mary University of London
Mile End Road
London E1 4NS

Queen Mary Research Ethics Committee
Hazel Covill
Research Ethics Administrator
Tel: +44 (0) 20 7882 7915
Email: h.covill@qmul.ac.uk

c/o Dr Hatice Gunes
Eng 211
Department of Electronic Engineering
Queen Mary University of London
Mile End Road
London E1 4NS

9th June 2013

To Whom It May Concern:

Re: QMREC2013/48 – Creative Vibes: Investigating Affect During Co-present Creative Interactions.

The above study was conditionally approved by The Queen Mary Research Ethics Committee (Review Panel F) on the 26th June 2013; full approval was ratified by delegated member's action on the 3rd July 2013.

This approval is valid for a period of two years, (if the study is not started before this date then the applicant will have to reapply to the Committee).

Yours faithfully

Ms Elizabeth Hall – QMREC Chair.

Patron: Her Majesty the Queen
Incorporated by Royal Charter as Queen Mary and Westfield College, University of London

A.2 Participant Forms

Queen Mary
University of London

Consent Form

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Study: **Creative Vibes: Investigating Affect During Co-present Creative Interactions**

Queen Mary Ethics of Research Committee Ref: **QMREC2013/48**

- Thank you for considering taking part in this research. The person organizing the research must explain the project to you before you agree to take part.
- If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.
- *I understand that if I decide at any other time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately.*
- *I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.*
- *I grant to the researcher and Queen Mary a non-exclusive, worldwide, irrevocable licence to use recordings of my performance for the purposes of research, including online postings (e.g. via You Tube) to publicise and illustrate the research results. Permission will be requested before the use of any video or images where you are identifiable. I retain all copyright and other intellectual property rights to my performance.*

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet (including Shimmer Informed Consent form) about the project, and understand what the research study involves.

Signed: Date:

Investigator's Statement:

I, Evan Morgan confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer.

Signed: Date:

Queen Mary
University of London

Information Sheet for Participants

Title of the Study - Creative Vibes: Investigating Affect During Co-present Creative Interactions

We would like to invite you to participate in this pilot research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.

In the event of you suffering any adverse effects as a consequence of your participation in this study, you will be compensated through Queen Mary University of London's 'No Fault Compensation Scheme'.

Background to the study

The term 'vibe' is often used to summarise the qualities of a co-present experience. Its definition is accordingly vague - *'A person's emotional state or the atmosphere of a place as communicated to and felt by others'* (Google Dictionary). This sense of shared emotion is particularly poignant when people are engaged in creative collaborations, such as an improvised jazz performance. We are interested in studying these situations in order to gain a more detailed understanding of what creates good or bad vibes and of the subsequent effect that they have on us.

Our research will utilise recent advances in emotion and behavioural sensing technologies (such as motion tracking, and physiological measurement) in an attempt to reveal and describe the subtleties of creative, co-present interactions.

From our experimental findings we intend to investigate the development of new collaborative musical interfaces that will facilitate or enhance emotional communication during co-present interactions. We believe that this is particularly important, since many contemporary music-related technologies focus on the individual and consequently act as a barrier to shared emotion and experience.

What your participation will involve

The experiment will last around 1 hour 45 minutes in total. You will receive a verbal description of the experiment, after which we will fit you with a number of sensors (details below). You will then enter the performance space with another participant. There will be two drums set up and a removable screen positioned between them (to allow or prevent you from viewing the other participant). You will complete a number of performances on the drums using one hand only. Afterwards we will ask you to watch a video of the improvised performances, and you will complete a couple of surveys to provide feedback on your experience. The exact format will be as follows:

1. Participant arrives, receives verbal description of experiment, reads information sheet and completes consent form (5 min).
2. Sensors fitted to participant (you can be instructed to do this alone in privacy) – the following sensors will be used: (20 min)
    - **ECG** – 4 sticky electrode pads placed on the chest and connected to a small wireless transmitter.
    - **GSR** – Two electrodes strapped to two fingers on the hand that the participant is not using for drumming. Wireless transmitter worn around wrist.
    - **EMG** – Three electrodes attached to inside of unused arm. Wireless transmitter worn around wrist.
    - **Eye Tracking** – This just involves wearing a pair of instrumented glasses, fitted with two small cameras.
    - **EEG** – Headset simply placed on head and electrodes positioned against scalp.
3. Two participants enter performance space and are instructed to carry out the following tasks with the screen *down* (20 min):
    - Hit the drum in time with a metronome for 1 minute, simultaneously.
    - Play a simple, pre-written piece for 5 minutes, together.
    - Improvise freely for 10 minutes, together.
4. Following a 5 minute break, participants are instructed to carry out the following tasks with the screen *up* (25 min):

  - Hit the drum in time with a metronome for 1 minute, one after the other.
  - Hit the drum in time with a metronome for 1 minute, both at the same time
  - Play a simple, pre-written piece for 5 minutes, together.
  - Improvise freely for 10 minutes, together.
5. The performance session concludes and the participants remove the sensors. They are now asked to watch video recordings of their two improvised performances and answer a series of questions for each 2 minutes of video (30 min).

Criteria for participation

In order to participate in this study you must be an experienced percussionist, with at least 2 years of experience performing live. You must be aged 20 or over and you must not have any know hearing or visual impairments. Additionally, due to the nature of the sensors you must not wear glasses.

Data Collection and Confidentiality

All of the data collected for this study will be held in accordance with the Data Protection Act 1988 and College regulations. We will not share your personal data with any third parties. The data you provide will be stored anonymously by giving the files code names that do not reveal your identity. We will be taking video and audio recordings of your performances, to which you retain the copyright. However, by agreeing to participate in the study you grant Queen Mary University of London permission (a licence) to use any material in the recording in which you own rights for the following non-commercial purposes: Research related investigation, public presentations of research work; still images may be used in written publications and research posters. Images containing identifiable features (such as your face or any tattoos) will not be used alongside your physiological data such that the two can be linked.

Risks

The sensors that we are using in this study are generally very safe, however there are some risks associated with unlikely events such as device failure. These risks are listed at the end of this form. The most likely risk is that your skin will react badly to the stick-on electrodes, although these are the same electrodes that are commonly used in hospitals. In the event of any allergic reaction the study will be stopped.

Financial Incentive

In order to encourage you to attend the study, and to account for the cost of your time and travelling expenses, we will pay you **£20** upon completion of the session. If you opt of the study before completion then we will not be able to provide any financial re-imbursement.

If you require any further information then please don't hesitate to contact us by emailing evan.morgan@eecs.qmul.ac.uk or calling 07815 481 041.

If this is unsuccessful, or not appropriate, please contact the Secretary at the Queen Mary Ethics of Research Committee, Room W117, Queen's Building, Mile End Campus, Mile End Road, London or research-ethics@qmul.ac.uk.

Shimmer Informed Consent Statement

As is the nature of a research prototype, Shimmer is an experimental device and consequently not fully tested to commercial product nor medical class regulatory approved standards. Consequently, although great care was taken in the design of this device, there is some inherent risk both with the design and manufacturing that you assume when the device is in close proximity to your body or the body of your test subjects:

These Risks Include:

There is a risk of electrical shock due to manufacturing defects or improper use (see usage guidelines and warnings).

There is also a risk of sustaining a burn due to a catastrophic failure of the device which could result from overheating of components.

There is a risk of radio interference with the operation of other electronic devices and we make no claims to the consequences of this.

There is a risk of some minor skin irritation from electrode pads over prolonged periods of time which may cause discomfort.

The device is not designed with proper safeguards for proper defibrillation. As such, electrodes must be removed before defibrillation is attempted.

Data privacy limitations: It should be understood from the outset and you should communicate to test subjects that the physiological data that is streamed, stored, and analysed through use of the device is not anonymized or privacy-protected in any way and you should take appropriate precautions in the protection and handling of such data in your research activities. Shimmer itself may buffer raw physiological data unencrypted on the integrated flash memory device.

Physiological data generated through use of Shimmer may indicate conditions that your test subject was previously unaware of prior to participation in research using the device.

There may be a risk of exposure to minute amounts of chemicals from the manufacturing process or the components themselves (such as latex, lead etc.).

There may be an increased risk of physical injury by the physical presence of the device on a test subject's body and you fully assume this responsibility.

Realtime Technologies Ltd is not liable for damage or loss of data when using the MicroSD card feature.

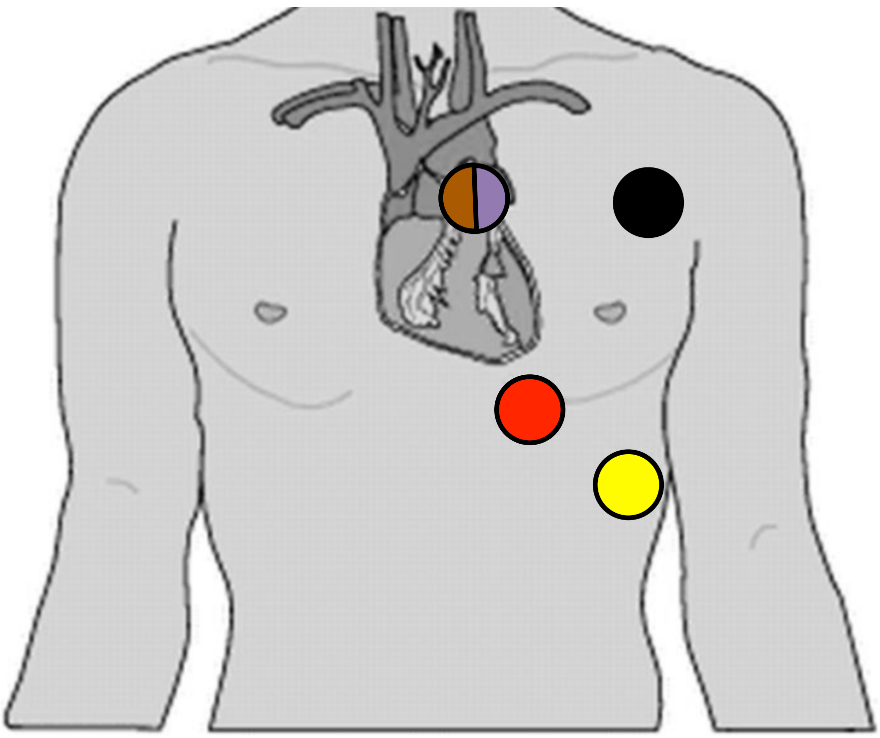
Some Shimmer peripherals rely on 3rd party driver support. While

Realtime/Shimmer Research have tested features using a typical system, in some cases the end user will need to contact the peripheral vendor for resolution of installation, compatibility, or operational issues.

By your use of Shimmer you acknowledge these and other risks inherent in the use of an experimental device and you assume full responsibility for testing this device with human subjects.

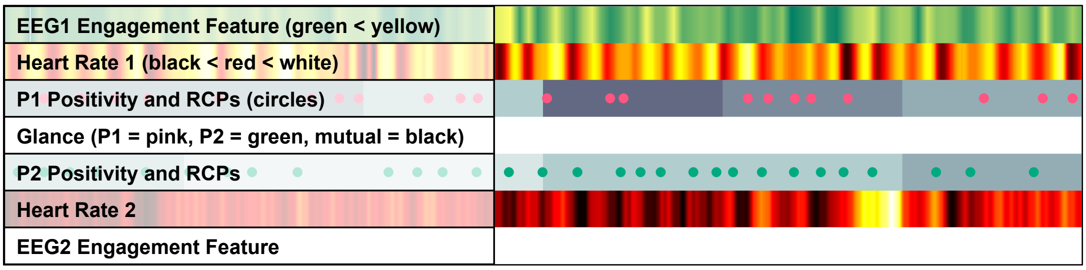
ECG Electrode Placement Guide

Please peel the plastic covers off the electrodes and stick them to your skin in the positions shown in the image below:

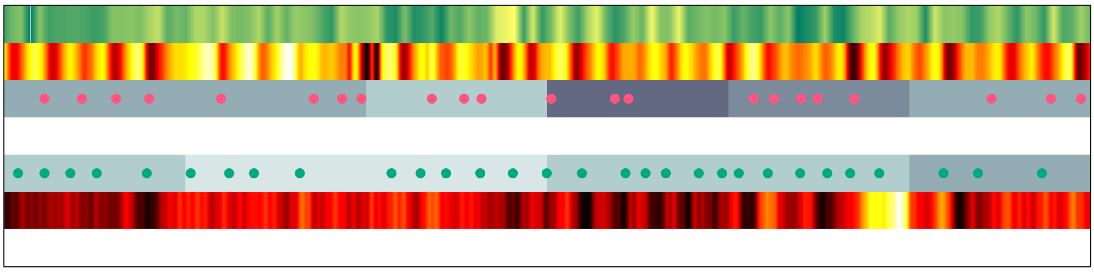


A.3 Strip Plot Visualisations

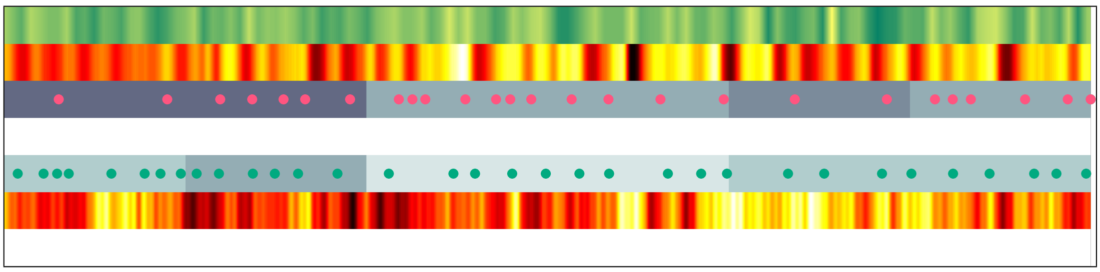
Strip plots were created, showing EEG (engagement feature), heart rate, positivity, rhythmic change points, and gaze (mutual and individual), for each pilot session and each participant. P1 = participant 1; P2 = participant 2. Lighter colours indicate higher readings, white indicates no reliable data collected. A strip plot example with labelled strips is shown below:



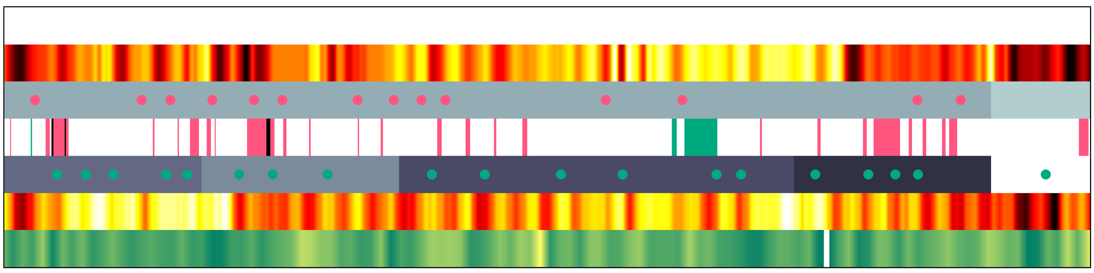
Strip plot for Group 1 in the visual condition:



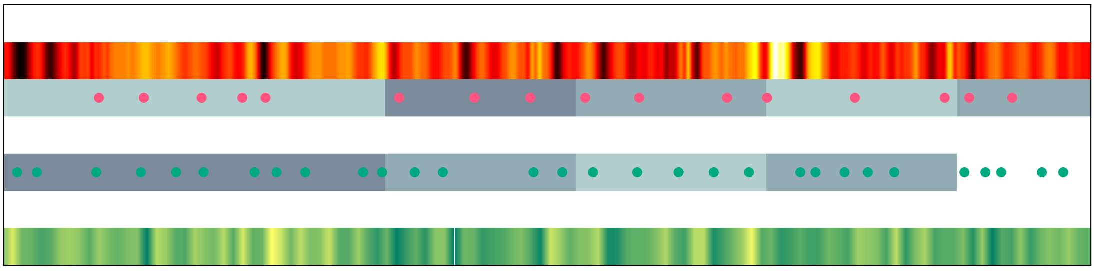
Strip plot for Group 1 in the non-visual condition:



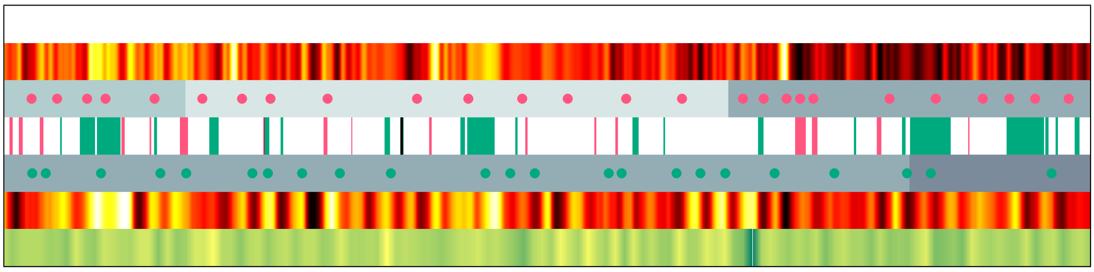
Strip plot for Group 2 in the visual condition:



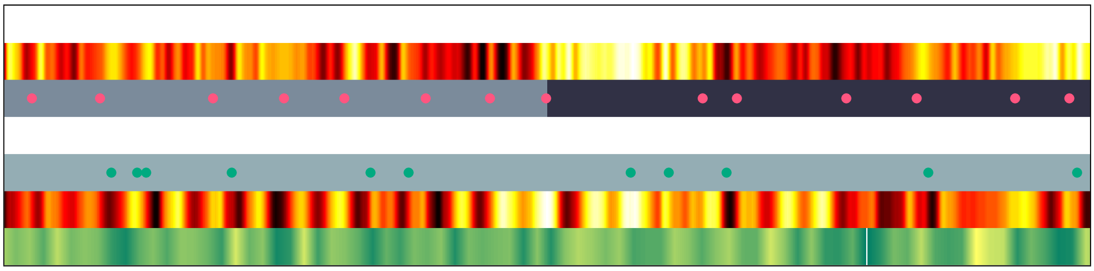
Strip plot for Group 2 in the non-visual condition:



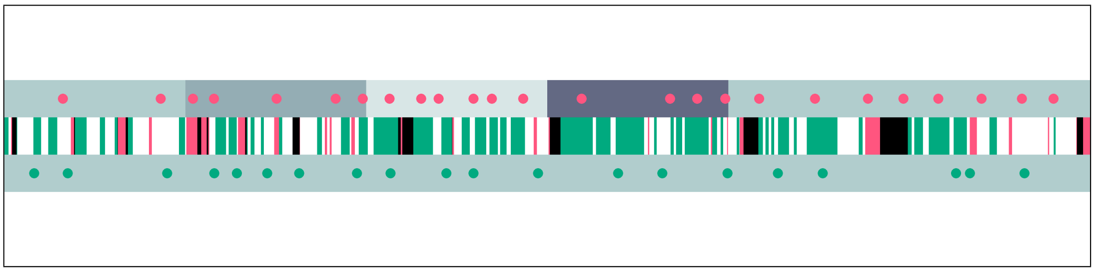
Strip plot for Group 3 in the visual condition:



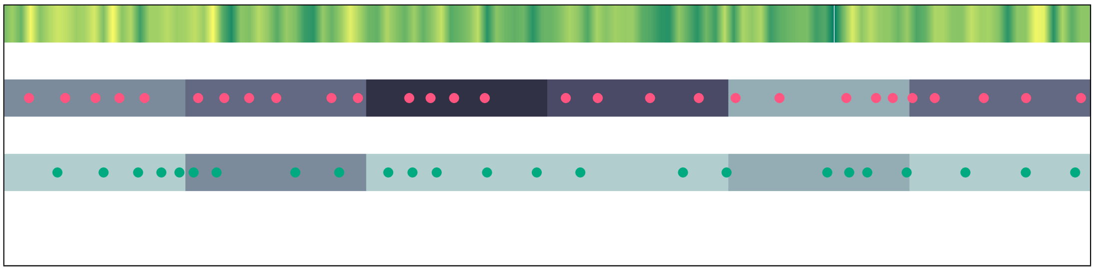
Strip plot for Group 3 in the non-visual condition:



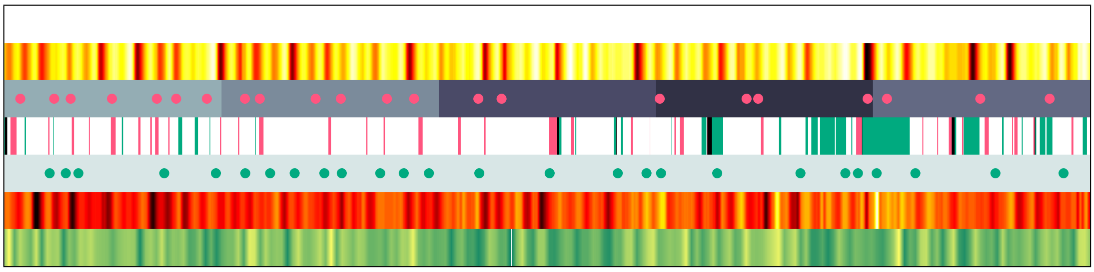
Strip plot for Group 4 in the visual condition:



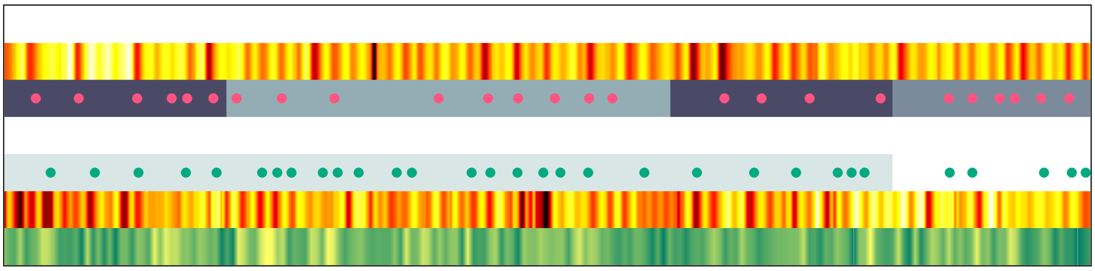
Strip plot for Group 4 in the non-visual condition:



Strip plot for Group 5 in the visual condition:



Strip plot for Group 5 in the non-visual condition:



A.4 SPSS Linear Mixed Model Command Syntaxes

Command syntax for the analysis of correlations between continuous features:

```
MIXED MeanInterBeatLag WITH NumMutualGlancemin
 /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001)
HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
 /FIXED=NumMutualGlancemin | SSTYPE(3)
 /METHOD=REML
 /PRINT=SOLUTION
 /RANDOM=INTERCEPT NumMutualGlancemin | SUBJECT(Par) COVTYPE(VC).
```

Command syntax for the analysis of correlations between continuous features and self-report scores across both conditions:

```
MIXED MeanBody BY Vis WITH Creativity
 /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001)
HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
 /FIXED=Creativity | SSTYPE(3)
 /METHOD=REML
 /PRINT=SOLUTION
 /RANDOM=INTERCEPT Creativity(Vis) | SUBJECT(Par) COVTYPE(VC).
```

Command syntax for the analysis of correlations between continuous features and self-report scores within a single condition:

```
MIXED Lbeta WITH Positivity
 /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001)
HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
 /FIXED=Positivity | SSTYPE(3)
 /METHOD=REML
 /PRINT=SOLUTION
 /RANDOM=INTERCEPT Positivity | SUBJECT(Par) COVTYPE(VC).
```

Command syntax for the analysis of effects of visibility condition:

```
MIXED Creativity BY Vis
 /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001)
HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
 /FIXED=Vis | SSTYPE(3)
 /METHOD=REML
 /PRINT=SOLUTION
 /RANDOM=INTERCEPT Vis | SUBJECT(Par) COVTYPE(VC).
```

Appendix B

Study 2 Materials

B.1 Ethical Approval

Queen Mary
University of London

Queen Mary, University of London
Room W117
Queen's Building
Queen Mary University of London
Mile End Road
London E1 4NS

Queen Mary Research Ethics Committee
Hazel Covill
Research Ethics Administrator
Tel: +44 (0) 20 7882 7915
Email: h.covill@qmul.ac.uk

c/o Dr Hatice Gunes
Eng 211
Department of Electronic Engineering
Queen Mary University of London
Mile End Road
London E1 4NS

2nd October 2014

To Whom It May Concern:

Re: QMREC2013/48 – Creative Vibes: Investigating Affect During Co-present Creative Interactions.

The above study was conditionally approved by The Queen Mary Research Ethics Committee (Review Panel F) on the 26th June 2013; full approval was ratified by delegated member's action on the 3rd July 2013.

A protocol amendment (reduced measurement taking; small changes to stimulus materials) was approved via QMERC Chair's Action on the 2nd October 2014.

This approval is valid for a period of two years, (if the study is not started before this date then the applicant will have to reapply to the Committee).

Yours faithfully

Ms Elizabeth Hall – QMREC Chair.

Patron: Her Majesty the Queen
Incorporated by Royal Charter as Queen Mary and Westfield College, University of London

B.2 Participant Forms

Information Sheet for Participants

Title of the Study - Creative Vibes: Investigating Affect During Co-present Creative Interactions

We would like to invite you to participate in this research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.

In the event of you suffering any adverse effects as a consequence of your participation in this study, you will be compensated through Queen Mary University of London's 'No Fault Compensation Scheme'.

Background to the study

The term 'vibe' is often used to summarise the qualities of a co-present experience. Its definition is accordingly vague - *'A person's emotional state or the atmosphere of a place as communicated to and felt by others'* (Google Dictionary). This sense of shared emotion is particularly poignant when people are engaged in creative collaborations, such as an improvised musical performance. We are interested in studying these situations in order to gain a more detailed understanding of what creates good or bad vibes and of the subsequent effect that they have on us.

Our research utilises recent advances in emotion and behavioural sensing technologies (such as eye tracking, and physiological measurement) in an attempt to reveal and describe the subtleties of creative, co-present interactions. We have also developed a simple device, the **LuminUs** (see Figure 1), which uses a coloured light display to provide the musician with feedback about the behaviour of their partner. Using eye-tracking glasses, the LuminUs can inform you when the other participant is looking towards you. Alternatively, it can inform you about the motion of the other participant, using a small accelerometer worn around the waist. We are interested in whether such devices can influence the interaction between musicians.

What your participation will involve

The experiment will last around 90 minutes in total. You will receive a verbal description of the experiment, after which we will fit you with a number of sensors (details below). You will then enter the performance space with another participant. There will be a set of electronic drum pads and a single keyboard (see Figure 2 on the next page). You will complete a number of performances where you will be asked to improvise an accompaniment to a short 2 minute video. After each performance you will complete a survey to provide feedback on your experience. The exact format will be as follows:

1. Participant arrives, receives verbal description of experiment, reads information sheet and completes consent form (10 min).
2. Sensors fitted to participant (you can be instructed to do this alone in privacy) – the following sensors will be used: (15 min)
    a **ECG** – 4 sticky electrode pads placed on the chest and connected to a small wireless transmitter worn around the waist.
    b **Eye Tracking** – This just involves wearing a pair of instrumented glasses, fitted with two small cameras.
3. Participants have a practice with the instruments before the study begins. (5 min)
4. Two participants enter performance space and are instructed to carry out the following baseline tasks:
    a Watch a short animation twice while baseline measurements are taken. (4 min)
    b Play along to the animation twice while baseline measurements are taken (4 min)
5. The study begins, participants are instructed to:
    a Improvise a live accompaniment to the animation twice (4 min)
    b Answer a series of questions about that session (2 min)
    c Repeat 5a and 5b a further 6 times (36 min)
6. The study ends. Participants remove the sensors and have a short de-briefing chat with the experimenter. Payment is given in cash. (10 min)

Total: 90 min

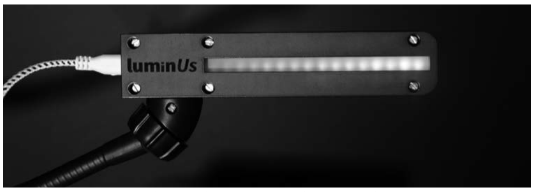


Figure 1 - The LuminUs device

Figure 2 - Experimental Set Up

Criteria for participation

In order to participate in this study you must be an experienced pianist or percussionist, with at least 5 years of experience performing live. You must be aged 18 or over and you must not have any known hearing or visual impairments. Additionally, due to the nature of the sensors you must not wear glasses.

Data Collection and Confidentiality

All of the data collected for this study will be held in accordance with the Data Protection Act 1988 and College regulations. We will not share your personal data with any third parties. The data you provide will be stored anonymously by giving the files code names that do not reveal your identity. We will be taking video and audio recordings of your performances, to which you retain the copyright. However, by agreeing to participate in the study you grant Queen Mary University of London permission (a licence) to use any material in the recording in which you own rights for the following non-commercial purposes: Research related investigation, public presentations of research work; still images may be used in written publications and research posters. Images containing identifiable features (such as your face or any tattoos) will not be used alongside your physiological data such that the two can be linked.

Risks

The sensors that we are using in this study are generally very safe, however there are some risks associated with unlikely events such as device failure. These risks are listed at the end of this form. The most likely risk is that your skin will react badly to the stick-on electrodes, although these are the same electrodes that are commonly used in hospitals. In the event of any allergic reaction the study will be stopped.

Financial Incentive

In order to encourage you to attend the study, and to account for the cost of your time and travelling expenses, we will pay you **£20** upon completion of the session. If you opt of the study before completion then we will not be able to provide any financial re-imbursement.

If you require any further information then please don't hesitate to contact us by emailing e.l.morgan@qmul.ac.uk or calling 07815 481 041.

If this is unsuccessful, or not appropriate, please contact the Secretary at the Queen Mary Ethics of Research Committee, Room W117, Queen's Building, Mile End Campus, Mile End Road, London or research-ethics@qmul.ac.uk.

Shimmer Informed Consent Statement

As is the nature of a research prototype, Shimmer is an experimental device and consequently not fully tested to commercial product nor medical class regulatory approved standards. Consequently, although great care was taken in the design of this device, there is some inherent risk both with the design and manufacturing that you assume when the device is in close proximity to your body or the body of your test subjects:

These Risks Include:

There is a risk of electrical shock due to manufacturing defects or improper use (see usage guidelines and warnings).

There is also a risk of sustaining a burn due to a catastrophic failure of the device which could result from overheating of components.

There is a risk of radio interference with the operation of other electronic devices and we make no claims to the consequences of this.

There is a risk of some minor skin irritation from electrode pads over prolonged periods of time which may cause discomfort.

The device is not designed with proper safeguards for proper defibrillation. As such, electrodes must be removed before defibrillation is attempted.

Data privacy limitations: It should be understood from the outset and you should communicate to test subjects that the physiological data that is streamed, stored, and analysed through use of the device is not anonymized or privacy-protected in any way and you should take appropriate precautions in the protection and handling of such data in your research activities. Shimmer itself may buffer raw physiological data unencrypted on the integrated flash memory device.

Physiological data generated through use of Shimmer may indicate conditions that your test subject was previously unaware of prior to participation in research using the device.

There may be a risk of exposure to minute amounts of chemicals from the manufacturing process or the components themselves (such as latex, lead etc.).

There may be an increased risk of physical injury by the physical presence of the device on a test subject's body and you fully assume this responsibility.

Realtime Technologies Ltd is not liable for damage or loss of data when using the MicroSD card feature.

Some Shimmer peripherals rely on 3rd party driver support. While Realtime/Shimmer Research have tested features using a typical system, in some cases the end user will need to contact the peripheral vendor for resolution of installation, compatibility, or operational issues.

By your use of Shimmer you acknowledge these and other risks inherent in the use of an experimental device and you assume full responsibility for testing this device with human subjects.

B.3 Post-performance Survey

Technology Testing Survey

Experiment Details

Page description:

1. Experiment ID *

2. Participant *

- 1
- 2

Participant Details

Page description:

3. Age *

years

4. Gender *

- Female
- Male

5. Please enter the name of your music college *

6. Please select your level of study *

- Undergraduate
- Postgraduate
- Diploma
- AS/A Level
- Other

7. For how many years have you been playing piano/percussion? *

years

Session 1

Page description:

Session ID (to be completed by the experimenter) *

- 0
- 1
- 2
- 3
- 4
- 5
- 6

Session 1

Page description:

Please answer the following questions based upon your experience of *each of the two attempts* in the last session.

On a scale of 0 to 10, how creative was each attempt? *

	Not at all Creative	Very Creative
Attempt 1		
Attempt 2		

Which attempt do you think produced the best accompaniment to the animation? *

Attempt 1 — Both were equal — Attempt 2

Session 1

Page description:

Regarding *the entire* last session (both attempts), how much do you agree/disagree with the following statements?

The other musician was leading the performance *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

I enjoyed this session *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

The other musician and I performed well as a pair *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

I was satisfied with my musical contributions *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

I felt engaged with the other musician *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

The session was boring *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

The other musician ignored my contributions *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

I liked the music we created *

Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

I felt connected to the other musician *

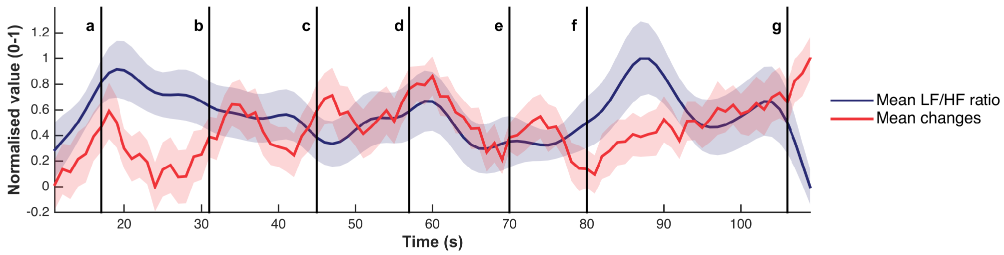
Strongly disagree	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
○	○	○	○	○	○	○

Page description:

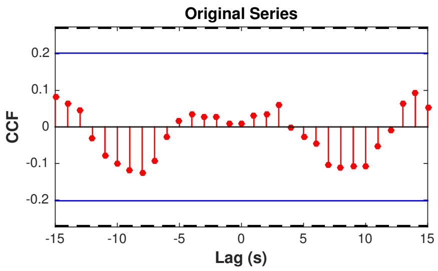
Thanks for your responses! Please return to your instrument ready for the next session.

Session 2

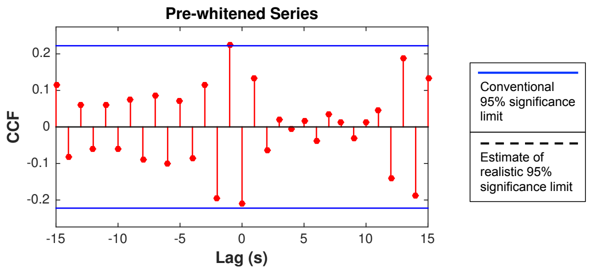
B.4 Time-series Analyses Plots



(a)

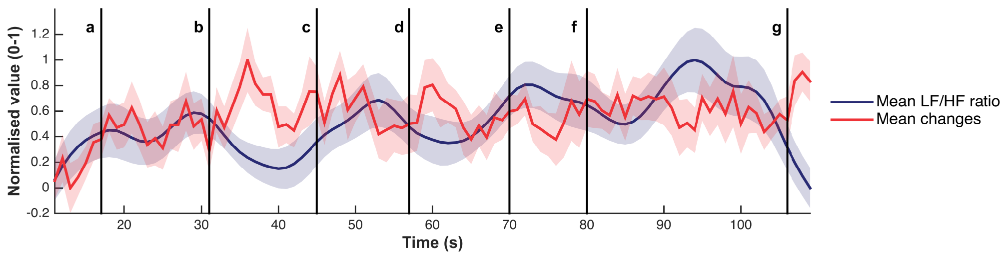


(b)

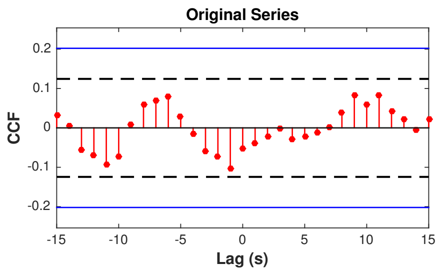


(c)

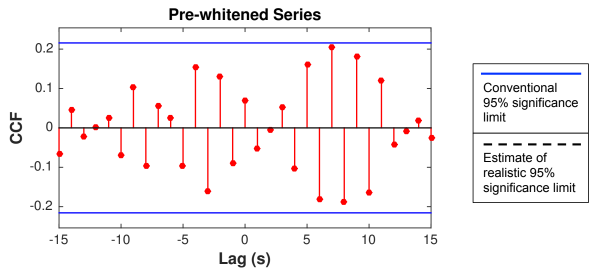
Figure B.4.1: Analysis of time series correlations between musical changes and LF/HF ratio over all performances for percussionists (P1) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating ± 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.



(a)

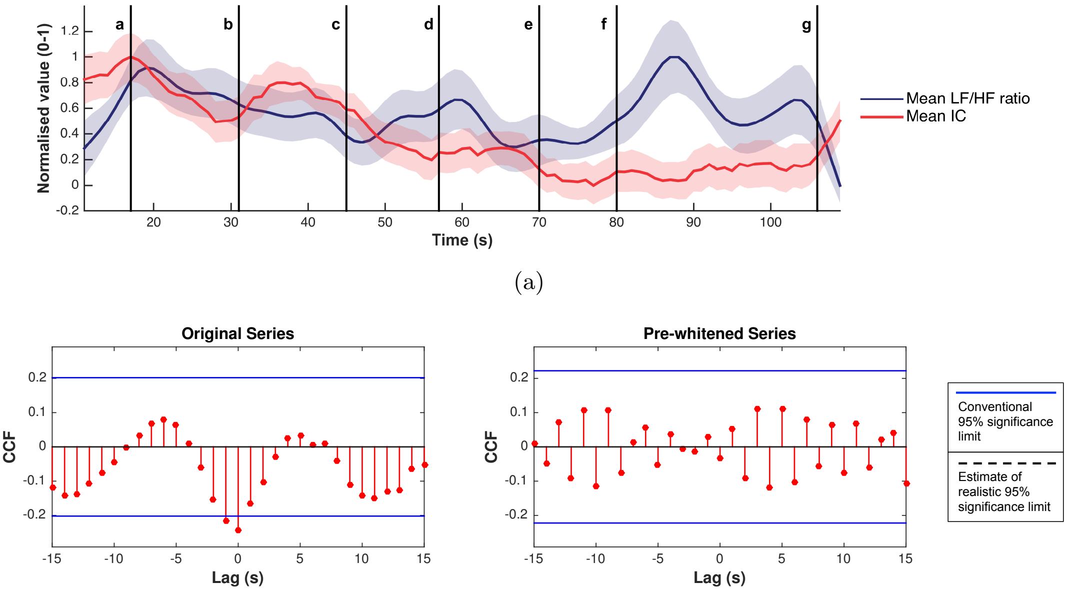


(b)



(c)

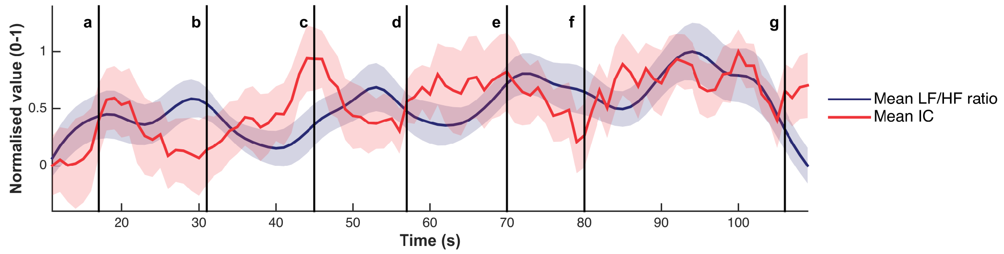
Figure B.4.2: Analysis of time series correlations between musical changes and LF/HF ratio over all performances for pianists (P2) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating $\pm$ 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.



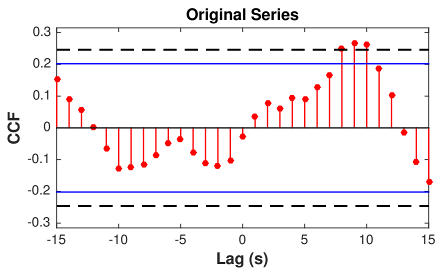
(a)

(b) (c)

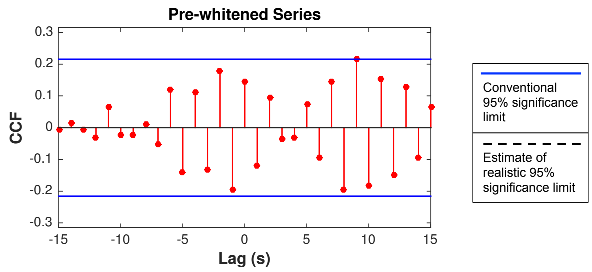
Figure B.4.3: Analysis of time series correlations between information content (IC) and LF/HF ratio over all performances for percussionists (P1) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating ± 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.



(a)



(b)



(c)

Figure B.4.4: Analysis of time series correlations between information content (IC) and LF/HF ratio over all performances for pianists (P2) only. Plot (a) shows the two averaged and normalised (0-1) time series, with the shaded areas indicating $\pm$ 1 SE. Plot (b) shows the sample cross-correlation function (CCF) for the two original series. Plot (c) shows the CCF for the pre-whitened series.

Appendix C

Study 3 Materials

C.1 Ethical Approval

Queen Mary
University of London

Queen Mary, University of London
Room W117
Queen's Building
Queen Mary University of London
Mile End Road
London E1 4NS

Queen Mary Ethics of Research Committee
Hazel Covill
Research Ethics Administrator
Tel: +44 (0) 20 7882 7915
Email: h.covill@qmul.ac.uk

c/o Dr Hatice Gunes
Eng 211
Department of Computer Science
Queen Mary University of London
Mile End Road
London

3rd September 2015

To Whom It May Concern:

Re: QMREC1547a – Creative Vibes: Investigating Affect During Co-present Creative Interactions.

I can confirm that Mr Evan Morgan has completed a Research Ethics Questionnaire with regard to the above research.

The result of which was the conclusion that his proposed work does not present any ethical concerns; is extremely low risk; and thus does not require the scrutiny of the full Research Ethics Committee.

Yours faithfully

Ms Hazel Covill – QMERC Administrator

C.2 Participant Forms

Queen Mary
University of London

Consent Form

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Study: **Creative Vibes: Investigating Affect During Co-present Creative Interactions**

Queen Mary Ethics of Research Committee Ref: **QMREC1547a**

- Thank you for considering taking part in this research. The person organizing the research must explain the project to you before you agree to take part.
- If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.
- *I understand that if I decide at any other time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately.*
- *I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.*
- *I grant to the researcher and Queen Mary a non-exclusive, worldwide, irrevocable licence to use recordings of my performance for the purposes of research, including online postings (e.g. via You Tube) to publicise and illustrate the research results. Permission will be requested before the use of any video or images where you are identifiable. I retain all copyright and other intellectual property rights to my performance.*

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: Date:

Investigator's Statement:

I, Evan Morgan confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer.

Signed: Date:

Queen Mary
University of London

Information Sheet for Participants

Title of the Study - Creative Vibes: Investigating Affect During Co-present Creative Interactions

We would like to invite you to participate in this research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.

In the event of you suffering any adverse effects as a consequence of your participation in this study, you will be compensated through Queen Mary University of London's 'No Fault Compensation Scheme'.

Background to the study

The term 'vibe' is often used to summarise the qualities of a co-present experience. Its definition is accordingly vague - *'A person's emotional state or the atmosphere of a place as communicated to and felt by others'* (Google Dictionary). This sense of shared emotion is particularly poignant when people are engaged in creative collaborations, such as an improvised musical performance. We are interested in studying these situations in order to gain a more detailed understanding of what creates good or bad vibes and of the subsequent effect that they have on us.

Our research utilises recent advances in emotion and behavioural sensing technologies in an attempt to reveal and describe the subtleties of creative, co-present interactions. We have also developed a simple device, the **LuminUs** (see Figure 1), which uses a coloured light display to provide the musician with feedback about the behaviour of their partner. Using eye-tracking glasses, the LuminUs can inform you when the other participant is looking towards you. Alternatively, it can inform you about the motion of the other participant, using a small accelerometer worn around the waist. We are interested in whether such devices can influence the interaction between musicians.

What your participation will involve

You will be asked to attend four sessions on separate days, each lasting around 90 minutes. All of the sessions will be held at Queen Mary University in the performance lab. The format for the sessions will be as follows:

Session 1: Rehearse a piece of music of your choice (1hr), followed by an informal interview (20 min)
Session 2: Compose a short (5-10min) piece (1hr), followed by an informal interview (20 min)
Session 3: Continue to compose/rehearse the piece (1hr), followed by an informal interview (20 min)
Session 4: Perform your piece in front of other participants (1hr), followed by a group discussion (30 min)

In each session you will be using the LuminUs. We will explain how it works, and show you how to set it up and use it. You'll have some flexibility in how you use it. Specifically, you will be able to choose which type of feedback (motion or gaze) you receive, and you can pick the positioning of the device and the eye tracking markers.

After each session there will be a short, informal interview, where we will discuss your use of the LuminUs. For the final session this will be a group discussion.

You can bring your own instruments, or use the ones we have in the lab (drum kit, guitars, keyboards).

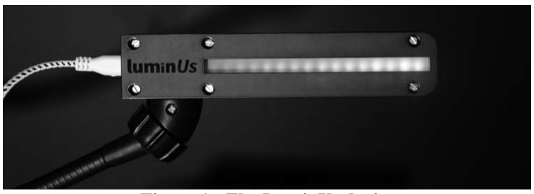


Figure 1 - The LuminUs device

Criteria for participation

In order to participate in this study you must be an experienced musician, with at least 5 years of experience performing live. You must be aged 18 or over and you must not have any known hearing or visual impairments. Additionally, due to the nature of the sensors you must not wear glasses.

Data Collection and Confidentiality

We will be taking video and audio recordings of your performances, and audio recordings of the interviews. All of the data collected for this study will be held in accordance with the Data Protection Act 1988 and College regulations. We will not share your personal data with any third parties. The data you provide will be stored anonymously by giving the files code names that do not reveal your identity. You retain the copyright to the recordings of your performances. However, by agreeing to participate in the study you grant Queen Mary University of London permission (a licence) to use any material in the recording in which you own rights for the following non-commercial purposes: Research related investigation, public presentations of research work; still images may be used in written publications and research posters. Images containing identifiable features (such as your face or any tattoos) will not be used alongside your data such that the two can be linked.

Financial Incentive

In order to encourage you to attend the study, and to account for the cost of your time and travelling expenses, we will pay you **£80** upon completion of the four sessions. If you opt of the study before completion then we will not be able to provide any financial re-imbursement.

If you require any further information then please don't hesitate to contact us by emailing e.l.morgan@qmul.ac.uk or calling 07815 481 041.

If this is unsuccessful, or not appropriate, please contact the Secretary at the Queen Mary Ethics of Research Committee, Room W117, Queen's Building, Mile End Campus, Mile End Road, London or research-ethics@qmul.ac.uk.

C.3 Thematic Analysis Codes

Name	Number Of Sources Coded	Number Of Coding References	Parent Node Name
Abilities	0	0	Group Attributes
Accelerometer on the leg	2	2	Positioning equipment
Accidental signalling	1	1	Confusion
Accuracy	2	2	Function
Aesthetics	3	3	Form
Alarm	3	7	Thinking and discussing
Ambiguity	1	1	Barriers and issues
An extra sense	1	1	Thinking and discussing
Appropriating	20	59	Usage
Attempted use	9	19	In practice
Audience impression	5	13	Expression and Communication
Awareness of other	5	11	Expression and Communication
Barriers and issues	1	1	Expression and Communication
Body language	3	3	Types of motion
Calibration	5	8	Practical issues
Changed approach	2	2	General influence
Clarity	1	1	Function
Colour meaning	1	5	Interpretation of feedback
Comfort	8	11	Form
Complexity	2	2	Practical issues
Composing	7	13	Situation
Composing influenced by tech	8	13	General influence
Confidence of musician	1	1	Abilities
Confusion	5	9	Problems
Definite motion made it spike	1	1	Interpretation of feedback
Design	0	0	Technology
Different situation	5	11	Uses for the LuminUs
Differentiating signals	1	9	Confusion
Discussing use	6	27	Thinking and discussing
Distraction	4	8	Problems
Doing a solo	9	15	Structure
Dynamics	11	25	Musical Affordances
Emotional expression	12	30	Expression and Communication
Excitement of seeing light go red	2	3	Response to feedback
Experimentation	5	6	In practice
Expression and Communication	0	0	Final Themes
Familiarity	3	3	Potential
Features of NVC	10	21	Expression and Communication
Form	0	0	Design
Function	0	0	Design
Functions of gaze	0	0	Expression and Communication
Gaze detection	6	6	Practical issues
General influence	1	1	Impact
Getting attention	13	24	Awareness of other
Getting used to it	2	3	Use over time
Glance conveying Instructions and questions	4	11	Functions of gaze
Glance length	5	7	Functions of gaze
Group Attributes	0	0	Final Themes
Hand motion	4	5	Types of motion
Head motion	8	12	Types of motion
Headset	1	1	Form
Hearing and listening	7	9	Expression and Communication
Imagined devices	2	7	Potential
Impact	0	0	Technology

Improved awareness	1	1	General influence
Improved efficiency	1	1	General influence
Improvisation	3	5	Musical style
In practice	1	1	Usage
In the studio	3	6	Situation
Instrument	9	26	Roles
Intensity	11	29	Musical Affordances
Interpretation of feedback	0	0	Impact
Just a 'feeling'	9	36	Emotional expression
Lag	3	3	Practical issues
Leading	7	8	Roles
Level of experience	1	1	Abilities
Light feedback	15	39	Function
Light show	2	10	Imagined devices
Lights are a quick reference	1	1	Response to feedback
Linking light to motion	1	3	Interpretation of feedback
Looking and lacking confidence	1	2	Functions of gaze
Looking at other to understand motion	1	1	Functions of gaze
Looking at other to understand tech	1	2	Response to feedback
Looking at other's LuminUs	1	1	Trust and confidence
Looking at other's motion for synchrony	1	1	Functions of gaze
Looking for appraisal	1	1	Functions of gaze
Looking to pass the ball	1	1	Functions of gaze
Lost in the music	1	4	Barriers and issues
Marker positioning	7	21	Positioning equipment
Meaning	13	68	Interpretation of feedback
Misunderstanding the tech	2	2	Confusion
More useful in larger group	3	5	Uses for the LuminUs
Motion and affect	6	13	Emotional expression
Motion as a count in	1	1	Appropriating
Motion information not useful	4	6	Not useful
Motion positioning	2	4	Positioning equipment
Motion sensitivity	1	1	Practical issues
Multitasking	7	7	Abilities
Musical Affordances	0	0	Final Themes
Musical Context	0	0	Final Themes
Musical style	11	32	Musical Context
Needing time to get used to it	1	1	Use over time
No use for tech	8	8	Thinking and discussing
Not useful	4	6	Problems
Originality	1	1	Function
Overcoming problems	3	3	Problems
Performance	1	3	Situation
Peripheral vision	7	14	In practice
Positioning equipment	0	0	In practice
Potential	0	0	Technology
Practical issues	0	0	Problems
Problems	0	0	Technology
Providing information	1	1	Functions of gaze
Range	2	3	Practical issues
Reading intentions	4	4	Expression and Communication
Relationships	7	17	Group Attributes
Response to feedback	8	27	Impact
Responsiveness of device	7	9	Function
Roles	12	30	Group Attributes
Rules	1	2	Thinking and discussing
Self awareness	3	6	General influence
Signalling and cueing	24	127	Expression and Communication

Signalling to sound engineer	1	2	Imagined devices
Simplicity	1	1	Function
Situation	19	110	Musical Context
Staring means not liking or wanting to stop	2	2	Functions of gaze
Structure	9	13	Musical Affordances
Subconscious influence of tech	2	3	General influence
Successful use	7	15	In practice
Tech influencing behaviour	5	10	Response to feedback
Tech positioning	14	52	Positioning equipment
Technology	0	0	Final Themes
Tempo	3	5	Musical Affordances
Testing	13	63	In practice
Thinking and discussing	1	1	Usage
Tonality	7	14	Musical Affordances
Trust and confidence	5	10	Problems
Trying to find a use	10	14	Thinking and discussing
Types of motion	0	0	Expression and Communication
Unnatural	5	9	Problems
Usage	0	0	Technology
Use over time	2	4	Usage
Useful for drummer	1	4	Uses for the LuminUs
Usefulness	1	3	Thinking and discussing
Uses for the LuminUs	14	46	Potential
Using to understand behaviour	6	16	Appropriating
Venue	5	6	Situation

C.4 Follow-up Survey

Name:

Age:

How many years have you been playing? …………………………………..

What style of music did you play this evening? …………………………………..

Did you end up using the technology in your rehearsals? YES / NO

If YES – how did you use it? If NO – why didn't you use it?

Did you use the technology differently during the performance? YES / NO

If YES – how did it differ?

What, if anything, did you like about the technology?

What, if anything, did you dislike about the technology?